AFOSR 70-0474TR

AD701792

January, 1970

Report ESL-R-412
M.I.T. Projects DSR 71526,
and 76265
AFOSR Grant 69-1724
NASA Grant NGL-22-009 (124)

1

ON THE OPTIMAL CONTROL OF LINEAR SYSTEMS WITH INCOMPLETE INFORMATION

Edison Tack-Shuen Tse

Electronic Systems Laboratory

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MASSACHUSETTS 02139

Department of Elecirical Engineering

294

AFOSR 70-0474TR

January, 1970 Report ESL-R-412

ON THE OPTIMAL CONTROL OF LINEAR SYSTEMS WITH INCOMPLETE INFORMATION

by

Edison Tack-Shuen Tse

This report is based on the unaltered thesis of Edison T-S Tse submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the Massachusetts Institute of Technology in January, 1970. This research was conducted at the Massachusetts Institute of Technology, Electronic Systems Laboratory, with support extended by the Air Force Office of Scientific Research under Grant AFOSR-69-1724 and by the National Aeronautics and Space Administration/Electronic Research Center under Grant NGL-22-009(124).

Electronic Systems Laboratory
Department of Electrical Engineering
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

ON THE OPTIMAL CONTROL OF LINEAR SYSTEMS
WITH INCOMPLETE INFORMATION

by

EDISON TACK-SHUEN TSE

S.B. and S.M., Massachusetts Institute of Technology
(1967)

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
November 1969

Signature of Author Edison T. S. Tse
Department of Electrical Engineering, November 5, 1969

Certified by Michael Athans
Thesis Supervisor

Accepted by ____________________
Chairman, Departmental Committee on Graduate Studies

ON THE OPTIMAL CONTROL OF LINEAR SYSTEMS
WITH INCOMPLETE INFORMATION

by

EDISON TACK-SHUEN TSE

Submitted to the Department of Electrical Engineering on November 5, 1969 in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

ABSTRACT

The control of linear systems with incomplete information is considered where the unknown disturbances and/or random parameters are assumed to satisfy some statistical laws.

The observer theory for linear systems is developed which generalizes the concepts due to Kalman and Luenberger pertaining to the design of linear systems which estimate the state f a linear plant on the basis of both noise-free and noisy measurements of the output variables. The Separation Theorem for linear system is then extended for such observers-estimators.

The problem of controlling a linear system with unknown gain is then considered. An open-loop-feedback-optimal control algorithm is developed which seems to be computationally feasible. Existence of such suboptimal control scheme is proved under the assumption that the uncertainties in the unknown gain will not grow in time. Convergence of such suboptimal control system to the truly optimal control system is considered. A computer program is developed to study the control of a variety thrid order systems with known poles but unknown zeroes. The experimental results serve to provide us with some more insights into the structure and behavior of the open-loop-feedback-optimal control systems.

Thesis Supervisor: Michael Athans

Title: Associate Professor of Electrical Engineering

ACKNOWLEDGEMENT

I would like to thank Professor Michael Athans in particular for the supervision of this research, his encouragement and support, his helpful suggestions and criticisms. Thanks are also extended to Professors Fred Schweppe and Ian Rhodes for their constructive criticisms and comments while acting as readers for this thesis. I would also like to express my sincere gratitude to Prof. Jan **Willems** for his interest and involvement in this research. Thanks are also due to Professors Leonard Gould and Roger Brockett who acted as my graduate Councelors during my graduate studies.

Many of my colleagues in the Control Theory Group of the Electronic Systems Laboratory have made contributions directly or indirectly to this research. In particular, I like to thank Dr. S. G. Greenberg, Dr. T. E. Fortmann, Messrs. K. Joseph, L. Kramer, and J. H. Davis. I am especially grateful to Professor A. H. Levis, Messrs. J. Gruhl and R. Moore.

I would like to thank Mr. P. Choy for typing most of the draft, Mrs. D. Orr for typing this thesis, and Messrs. H. Tonsing and A. Giordani of ESL Drafting Department for their help in the figures.

Finally, I thank my parents and my wife who provided a great deal of help in the form of love and encouragement.

The computations were carried out at the M.I.T. Information Processing Center.

This research was supported in part by the U.S. Air Force Office of Scientific Research under grant AFOSR 69-1742 and in part by the NASA Electronics Research Center under grant NASA NGL-22-009-(124).

CONTENTS

CONTENTS (Contd.)

CONTENTS (Contd.)

DISCLAIMER NOTICE

THIS DOCUMENT IS BEST QUALITY PRACTICABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

OR are
Blank pgs
that have
Been Removed

CHAPTER I

INTRODUCTION

In recent years, deterministic optimal control theory has come to its full maturity. Text books [57], [43] have been written which are devoted to the theory and application of modern control theory. In deterministic control theory, it is assumed that the future effect of any present control action is exactly known; this class of control problems is often called control with complete information [73]. In many situations, the necessity of control arises from the fact that there are perturbing disturbances and/or component failures in the physical system. These uncertain phenomena prevent us to deduce exactly the future effect of all present actions, and thus deterministic control theory may not be strictly applicable. The classes of control problems where future effect cannot be predicted exactly are called control with incomplete information. There are cases where the uncertain phenomena can be appropriately modelled as stochastic processes, so stochastic models and stochastic control theory can be applied [4], [74]. There are also cases where the chance phenomena have no statistical regularity, in these situations, the game-theoretic approach [65] to obtain min-max control may be more appropriate.

In this thesis we shall study some classes of problems with incomplete information. First we assume that the system being controlled is linear (either discrete time or continuous time). The disturbance and random parameters are assumed to satisfy some statistical laws. In the beginning, we assume that the only sources of uncertainty are the driving and/or observation disturbances. The statistical laws of disturbances are assumed to be

known. Then, we shall consider the case where some parameters of the system are unknown but satisfy some statistical laws.

In Chapter II, some mathematical preliminaries are collected for ease of references. Probability theory is treated briefly from a measure-theoretic approach. Facts about linear stochastic difference (and differential) equations, and stochastic optimization problems are included for the sake of completeness. The sections on Generalized Riccati Equations are new results and will be useful in later discussions. The theory for observers for discrete time and continuous time linear systems is developed in Chapters III and IV. The conceptual framework is that an observer is a device which will supply complementary information about all recoverable uncertainties. The observer theory is applied to estimation problems where we have only partial observation of the states in the presence of observation noise which may be degenerate or even totally singular. The results will include the Kalman filter [39], [50] as a special case.

In Chapter V, we consider the optimal control of linear system with known dynamics with respect to quadratic criterion. The uncertainties arise from driving and/or observation disturbances with known statistical laws. One such class of problems had been considered before by Joseph and Tou [56], Streibel [59] and Wonham [22]. They made the assumption that the observation noise is nondegenerate Gaussian white noise process (see Section 2.2). In this work, this assumption is relaxed. It is assumed that the observation noise may be: 1) nondegenerate Gaussian white noise, 2) degenerate Gaussian white noise, 3) colored observation noise, 4) totally singular (i.e., noise-free observations) or 5) the sum of colored and white

Gaussian noise. The approach follows that of Wonham's [22] and the technique is the dynamic programming method.

The control of linear systems with unknown gain parameters is considered in Chapter VI. The open-loop-feedback-optimal approach is used to derive a suboptimal control sequence which appears to be computationally feasible. The technique used is that of the matrix minimum principle. Analytical studies on the overall suboptimal control system are carried out and the asymptotic behavior of the overall suboptimal control system is derived. Computer simulations for some third order linear systems were carried out based on the theoretical results obtained in Chapter VI. The experimental results are discussed in Chapter VII. Conclusions and some topics for further research are listed in Chapter VIII.

The perspective and comparison of this work with published references are done at the end of each chapter. In this contribution, we develop the observer theory which provides a deeper understanding of the structure of state estimators in the case of nondegenerate, degenerate, singular, or colored observation noise. The theory unifies some seemingly different concepts of Kalman filter, Luenberger observer and exponential estimator, and treated them in one general framework. Then we have the extension of the Separation Theorem for such observers-estimators. Finally, we have developed the open-loop feedback optimal control algorithm for the linear stochastic systems with unknown constant or random gain parameters; theoretical and experimental studies are carried out to this class of problems which provide us with some insights into the structure and behavior of the overall control system.

Notations:

Lower case underscored letters stand for vectors (e.g., $\underline{x}$, $\underline{y}$); upper case underscored letters stand for matrices (e.g., $\underline{A}$, $\underline{B}$). Noise disturbances are denoted by lower case underscored Greek letters (e.g., $\underline{\xi}$, $\underline{\eta}$). Lower case letters with subscripts will denote components (e.g., x_i will be the i-th component of the vector $\underline{x}$, a_{ij} will be the ij-th element of matrix $\underline{A}$).

The transpose of a matrix $\underline{A}$ is denoted by $\underline{A}'$. The transpose of a column vector, $\underline{x}$, is a row vector and is denoted by $\underline{x}'$.

Let $\underline{A}$ be an nxn square matrix; the trace of $\underline{A}$ is defined as

$$\mathrm{tr}\,\underline{A} = \sum_{i=1}^{n} a_{ii} \quad .$$

Let $H(x_{11}, x_{12}, \ldots, x_{nm})$ be a scalar function; we shall denote it by $H(\underline{X})$. The gradient matrix is defined by

$$\frac{\partial H(\underline{X})}{\partial \underline{X}} = \left(\frac{\partial H(x_{11}, \ldots, x_{nm})}{\partial x_{ij}}\right) \quad .$$

M_{nm} will denote the set of all n×m matrices.

R_n will denote the product space of ordered n-tuples of real numbers, we shall denote the elements in R_n by column vector $\underline{x}$.

I will denote the set of all integers and $I_{[i,j]}$ will denote the set of integers $\{i, i+1, \ldots, j\}$, $i \le j$.

CHAPTER II

MATHEMATICAL PRELIMINARIES

2.1 Introduction

The purpose of this chapter is to introduce the mathematical results which will be used frequently in the later chapters. Some of these results are known in the literature while some are due mainly to the author.

In Section 2.2, probability theory is treated briefly using the measure-theoretic approach. Except for the precise basic definitions, the treatment is physical rather than mathematical. For a detailed and rigorous mathematical treatment, see Doob [1] and Loeve [2]. A rigorous mathematical consideration on conditional expectation and conditional distribution of a random vector is given. In the opinion of the author, a thorough understanding of these concepts is vital in most stochastic optimization problems.

In Section 2.3, linear stochastic difference and differential equations are treated to the extent that some of the discussions in later chapters will require for the sake of completeness.

In Section 2.4, the matrix minimum principle and optimality criteria are considered to some detail. The matrix minimum principle can allow us to deduce the necessary conditions for optimality for some special problems, whereas the optimality criteria provides us a test to see whether a certain solution is optimal.

In most control and filtering problems, we shall encounter a matrix Riccati Difference or Differential equation. To foresee there generalized matrix Riccati difference and differential equations are investigated in detail in sections 2.5 and 2.6. The results obtained in these sections are new, while the approach follows that of Wonham's [32].

From Section 2.2 to Section 2.4, the results are known. The discussions in these sections are by no means exhaustive: detailed references are given in Section 2.7 to indicate where more extensive results can be found.

2.2 Probability Theory

Definition 2.2.1: Let Ω be a set. A σ-algebra (Borel Field, σ-field) on Ω, F, is a class of subsets of Ω, such that it has the following properties:

a) $\Omega \varepsilon F$

b) If $A \varepsilon F$, then $A^c \varepsilon F$†

c) If $A_i \varepsilon F$, $i = 1, a, \ldots$, then

$$\bigcup_{i=1}^{\infty} A_i \varepsilon F \quad ; \quad \bigcap_{i=1}^{\infty} A_i \varepsilon F \quad .$$

The pair (Ω, F) consisting of a set Ω and a σ-field F of subsets of Ω is called a measurable space. The elements of Ω are called F-measurable sets, or just measurable sets if there is no ambiguity. In probability theory, the set Ω represents the sample space, and F represents the collections of possible events.

Definition 2.2.2: Let (Ω_1, F_1) and (Ω_2, F_2) be two measurable spaces. A mapping f of Ω_1 onto Ω_2 is said to be measurable if it satisfies the condition:

$$f^{-1}(A) \varepsilon F_1 \quad \text{for every } A \varepsilon F_2 \quad .$$

Definition 2.2.3: Let Ω be a set, and $(f_i)_{i\varepsilon I}$ a family of mappings of Ω into measurable spaces $(\Omega_i, F_i)_{i\varepsilon I}$. The σ-algebra generated by $(f_i)_{i\varepsilon I}$ is the smallest σ-algebra on Ω with respect to which all functions $(f_i)_{i\varepsilon I}$ are measurable, and is denoted by $F(f_i, i \varepsilon I)$.

---

† A^c denotes the complement of A.

From the above two definitions, we see that if F' is the σ-algebra generated by $(f_i)_{i\varepsilon I'}$, and F'' is the σ-algebra generated by $(f_i)_{i\varepsilon I''}$ then $F' \subseteq F''$ if and only if $I' \subseteq I''$ while I' and I'' are both countable.

In general, a basic measurable space (Ω, F) is assumed to be given which describes the underlying uncertainty of random phenomena. Such a measurable space is of an abstract nature; how the uncertainties reveal themselves depends on the type of experiments we perform to obtain observations, the outcomes of which we usually referred to as statistics. In abstract mathematical formulation, we let (Ω_1, F_1) be another measurable space, where we call Ω_1 the observation space and F_1 the collections of all possible observations. A measurable function, f, from Ω to Ω_1 is called the <u>observation statistic</u>. Let $\tilde{F} \subset F$ be a sub-σ-algebra, an observation statistic, f, is said to be $\tilde{F}$-measurable if $F(f) \subset \tilde{F}$. Special cases of observation statistics are random vectors $(\Omega_1 = R^n)$ and random processes (Ω_1 is the set of functions defined on $[o, T]$ with values in R^n).

<u>Definition 2.2.4</u>: Let (Ω, F) be a measurable space. A <u>probability law</u> on this space is an abstract positive measure μ defined on F,[†] and having $\mu(\Omega) = 1$. The triplet (Ω, F, μ) is called a probability space.

Let (Ω, F, μ) be a basic probability space, and let (Ω_1, F_1) be another measurable space representing the observations with a statistic f which maps Ω onto Ω_1. We can define a probability law on (Ω_1, F_1) by defining $\mu_f(A) = \mu(f^{-1}(A))$; $A \varepsilon F_1$. We shall call μ_f the statistical law of μ under f; this law is also called the law of distribution of the statistic f.

[†] $\mu(\cdot)$ is a set function defined on F with the property of countable additivity, i.e., if $A_n \varepsilon F$, $n \varepsilon I$, are disjoint, then we have

$$\mu\left(\bigcup_n A_n\right) = \sum_n \mu\left(A_n\right) \quad .$$

Definition 2.2.5: Let (Ω, F, μ) be a probability space, let F_1 be a sub-σ-field of F, and let $\underline{x}$ be an integrable real-valued random vector. A conditional expectation of $\underline{x}$ relative to F_1 is an integrable F_1-measurable real-valued random vector $\underline{y}$ such that

$$\int_A \underline{x}(w)d\mu(w) = \int_A \underline{y}(w)d\mu(w) \quad \text{for every } A \in F_1 \quad . \tag{2.2.1}$$

By the Radon-Nikodym theorem, such a random vector, $\underline{y}$, exists and is unique a.s. (almost surely): i.e., if $\underline{y}'$ is another random vector satisfying (2.2.1), then

$$\mu\{\omega: \underline{y}(\omega) = \underline{y}'(\omega)\} = 1 \quad . \tag{2.2.2}$$

For this reason, we may simply write such $\underline{y}$ as $E\{\underline{x}|F_1\}$. The conditional expectation of the indicator of $A \in F$ with respect to F_1, $E\{I_A|F_1\}$, is also called the conditional probability of A relative to F_1. Note that this "probability" is a random variable defined up to an a.s. equality, and not a number.

Lemma 2.2.6: Let (Ω, F, μ) be a probability space. Let F_1, F_2 be sub-σ-algebra of F with $F_1 \subset F_2$. Then

$$E\left\{E\{\underline{y}|F_2\}\middle| F_1\right\} = E\{\underline{y}|F_1\} \quad \text{a.s.} \tag{2.2.3}$$

where $\underline{y}$ is any μ-integrable real-valued random vector.

Proof: By definition 2.2.5, we have for all $A \in F_2$

$$\int_A E\{\underline{y}|F_2\}d\mu = \int_A \underline{y}\, d\mu \quad . \tag{2.2.4}$$

By assumption, $F_1 \subset F_2$, therefore (2.2.4) holds for all $A \in F_1$. Therefore

$$\int_A E\{E\{\underline{y}|F_2\}|F_1\}d\mu = \int_A E\{\underline{y}|F_2\}d\mu =$$

$$\int_A \underline{y}\, d\mu = \int_A E\{\underline{y}|F_1\}d\mu \quad ; \quad A \in F_1 \quad . \tag{2.2.5}$$

Now (2.2.4) follows from the a.s. uniqueness of (2.2.1).

Lemma 2.2.7: Let (Ω, F, μ) be a probability space. Let F_1 be a sub-σ-algebra of F. Let y be a μ-integrable random variable and x is a F_1-measurable random variable, then

$$E\{xy|F_1\} = x\, E\{y|F_1\} \quad . \tag{2.2.6}$$

Equation (2.2.6) is true when x is a simple function, and the general case follows using the approximation procedure. For a detailed discussion, see [1], [2].

Let f be an observation statistic on Ω; i.e., f is a measurable function from (Ω, F) onto (Ω_1, F_1). Let $F(f)$ be the σ-algebra generated by f. Such a statistic induces a conditional probability $E\{I_A|F(f)\}$on F. If there exists a function $P_f(A, \omega)$ such that for each $\omega \in \Omega$, $P_f(A, \omega)$ defines a probability measure on F and for fixed $A \in F$.

$$P_f(A, \omega) = E\{I_A|F(f)\} \qquad \text{a.s.} \tag{2.2.7}$$

then $P_f(A, \omega)$ is called a conditional measure on F relative to the statistics f. Unfortunately, such $P_f(A, \omega)$ may not exist, and so it may not always be possible to define a conditional measure on F relative to a certain statistic.[1] Let g be another statistic and $F(g)$ is the σ-algebra generated by g. If there exists a conditional measure defined on $F(g)$; i.e., if there is a function $P_f(A, \omega)$ such that for each $\omega\epsilon\Omega$, $P_f(A,\omega)$ defines a measure on $F(g)$,

and for fixed $A \in F(g)$

$$P_f(A, \omega) = E\left\{I_A | F(f)\right\} \quad \text{a.s.} \tag{2.2.8}$$

then one can define the law of distribution of g in the regular manner. Doob[1] had proved that if the statistic g is a random vector (say $\underline{y} \in R^n$) then the conditional measure on $F(\underline{y})$, $P_f(A, \omega)$, $A \in F(\underline{y})$, exists and so the conditional distribution of $\underline{y}$ is well defined (a.s.). Let us denote the conditional distribution by $P_f(\underline{Y}, \omega)$ which defines a conditional measure on the Borel set of R^n through $\underline{y}$. If $\Phi(\underline{y})$ is a measurable function of n-variables with values in R^m, then almost surely, we have:[1]

$$E\left\{\Phi(y) | F(f)\right\} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \Phi(\underline{y})\, P_f\left(dy_1 \ldots dy_n, \omega\right) \quad . \tag{2.2.9}$$

We can visualize F(f) as the σ-algebra which contains, in a loose sense, all the statistical information conveyed by the observation statistic about the total underlying uncertainty of the basic sample space. On the other hand, the conditional measure $P_f(A, \omega)$, $A \in F(\underline{y})$, describes the statistical information of f conveyed <u>about</u> the random vector $\underline{y}$. In view of this intuitive interpretation we have the following definition.

<u>Definition 2.2.8</u>: Let (Ω, F, μ) be a probability space, and let F_1, F_2 and F_3 be sub-σ-algebras of F. F_1 and F_3 are said to be <u>conditionally independent relative</u> to F_2 if for any random vectors, $\underline{y}_1$ which is F_1-measurable, and $\underline{y}_3$ which is F_3-measurable; we have

$$E\left\{\underline{y}_1\, \underline{y}_3' | F_2\right\} = E\left\{\underline{y}_1 | F_2\right\} E\left\{\underline{y}_3' | F_2\right\} \quad . \tag{2.2.10}$$

Let f_1, f_2 be two observation statistics; f_1 and f_2 are said to be <u>independent</u> if $F(f_1)$ and $F(f_2)$ are conditional independent relative to $F_2 = (\phi, \Omega)$, or we say $F(f_1)$ and $F(f_2)$ are <u>independent</u>.

Let (Ω, F, μ) be a basic probability space, and let $\underline{y}$ be a random n-vector on Ω; this induces a law of distribution on R^n through the statistic $\underline{y}$. Let $F_1 \subset F$ such that F_1 and $F(\underline{y})$ is independent relative to $F_2 = \{\phi, \Omega\}$. Then for arbitrary $B \in F_1$, we have for $A \in F(\underline{y})$, and $\underline{\Phi}(\underline{y})$ measurable in R^n:

$$\int_B E\{\underline{\Phi}(\underline{y})|F_1\} \, d\mu = \int_B \underline{\Phi}(\underline{y}) \, d\mu = \int_\Omega I_B \, \underline{\Phi}(\underline{y}) \, d\mu$$

$$= \int_B d\mu \int_\Omega \underline{\Phi}(\underline{y}) \, d\mu \tag{2.2.11}$$

$$= \int_B \left(\int_\Omega \underline{\Phi}(\underline{y}) \, d\mu\right) d\mu \qquad B \in F_1 \quad .$$

Therefore, we have

$$E\{\Phi(\underline{y})|F_1\} = \int_\Omega \Phi(\underline{y}) \, d\mu \qquad \text{a.s.} \quad . \tag{2.2.12}$$

In particular if $\underline{\Phi}(\underline{y}) = I_A$, $A \in F(\underline{y})$, then (2.2.11) and (2.2.12) become

$$\mu(A \cap B) = \mu(A) \, \mu(B) \qquad ; \qquad E\{I_A|F_1\} = \mu(A) \qquad \text{a.s.} \quad . \tag{2.2.13}$$

This implies that if F_1 and $F(\underline{y})$ are independent, the conditional distribution of $\underline{y}$ relative to F_1 is the same as the unconditional distribution of $\underline{y}$. Physically, this says that F_1 reveals no information about $\underline{y}$. In many cases, F_1 is generated by some observation statistics, $f_1, \ldots, f_n$; so if $\underline{y}$ is independent of $F_1 = F(f_i, i = 1, \ldots, n)$, this means that the observation of $f_1, \ldots, f_n$ reveals no statistical information about $\underline{y}$.

Let $\underline{x}$ be a random vector defined on the basic probability space (Ω, F, μ). $\underline{x}$ is called a Gaussian random vector if it has the distribution law.[3]

$$\mu_x(A) = \frac{1}{|2\pi\underline{\Sigma}|^{\frac{1}{2}}} \int_{x \in A} \exp - \frac{1}{2} (\underline{x} - \underline{m})' \, \underline{\Sigma}^{-1} (\underline{x} - \underline{m}) \, d\underline{x} \tag{2.2.14}$$

where

$$\underline{m} \triangleq E\{\underline{x}\} = \int_{\Omega} \underline{x}(\omega)\mu(d\omega) \quad ; \quad \underline{\Sigma} \triangleq E\{(\underline{x} - \underline{m})(\underline{x} - \underline{m})'\}$$

$$= \int_{\Omega} (\underline{x}(\omega) - \underline{m})(\underline{x}(\omega) - \underline{m})'\mu(d\omega) \quad . \tag{2.2.15}$$

$\underline{m}$ is called the mean or expectation of the random vector $\underline{x}$, and $\underline{\Sigma}$ is called the covariance matrix of the random vector $\underline{x}$. From (2.2.14), we see that the statistical law of a Gaussian random vector is specified completely by its mean and covariance. We shall always denote a Gaussian vector with mean $\underline{m}$ and covariance $\underline{\Sigma}$ by the symbol $\mathcal{G}(\underline{m}, \underline{\Sigma})$.

Two Gaussian vectors $\underline{x}_1$, $\underline{x}_2$ are independent if and only if[3]

$$E\{\underline{x}_1\ \underline{x}_2'\} \triangleq \int_{\Omega} \underline{x}_1(\omega)\underline{x}_2'(\omega)\mu(d\omega) = \int_{\Omega} \underline{x}_1(\omega)\mu(d\omega)\int_{\Omega} \underline{x}_2'(\omega)\mu(d\omega)$$

$$= E\{\underline{x}_1\} \cdot E\{\underline{x}_2'\} \quad . \tag{2.2.16}$$

Let $\underline{x}(t)$, $t \in [t_0, T]$, be a random n-vector process defined on the probability space (Ω, F, μ). $\underline{x}(t)$, $t \in [t_0, T]$, is called a Gaussian random n-vector process if for any finite set $\{t_1, \ldots, t_m\}$, $t_i \in [0, T]$ the vector

$$\underline{x}(\omega) = \begin{bmatrix} \underline{x}(t_1, \omega) \\ \vdots \\ \underline{x}(t_m, \omega) \end{bmatrix}$$

is a Gaussian random nm-vector.

Another observation statistic which we shall consider in the later chapters is the "Gaussian White Noise Process." Different interpretations of this kind of process are available. One may view it as a formal derivative of a Wiener Process,[4] or as a generalized random process[5] where the observation space is the set of linear functional on the class of test

functions. We shall not consider these interpretations in detail; no matter what interpretation one adapts, a Gaussian White Noise, $\underline{\xi}(t)$, has the following properties:

1) $\int_{t_0}^{t} \underline{\xi}(\tau)d\tau$ is Gaussian for all $t \in [t_0, T]$ with mean $\int_{t_0}^{t} \underline{m}(\tau)d\tau$ and covariance $\int_{t_0}^{t} \underline{R}(\tau)d\tau$, $R(\tau)$ is measurable and in L_2 locally.

2) $\int_{t_0}^{t_1} \underline{\xi}(\tau)d\tau, \int_{t_1}^{t_2} \underline{\xi}(\tau)d\tau \ldots \int_{t_{n-1}}^{t_n} \underline{\xi}(\tau)d\tau$, $t_0 < t_1 \ldots < t_n < T$, are independent.

Let F_t be the σ-algebra generated by $\underline{\xi}(\tau)$, $t_0 \le \tau \le t$, then (2.2.12) and the properties of Gaussian White Noise imply that

$$E\left\{\int_t^\sigma \underline{\xi}(\tau)d\tau \,\middle|\, F_t\right\} = \int_t^\sigma \underline{m}(\tau)d\tau \tag{2.2.17}$$

$$E\left\{\left(\int_t^\sigma \underline{\xi}(\tau)d\tau - \int_t^\sigma \underline{m}(\tau)d\tau\right)\left(\int_t^\sigma \underline{\xi}(\tau)d\tau - \int_t^\sigma \underline{m}(\tau)d\tau\right)' \,\middle|\, F_t\right\} = \int_t^\sigma \underline{R}(\tau)d\tau \quad . \tag{2.2.18}$$

2.3 Linear Stochastic Difference and Differential Equations

Consider a discrete-time linear system described by

$$\underline{x}(k+1) = \underline{A}(k)\,\underline{x}(k) + \underline{\xi}(k) \quad ; \quad k = 0, 1, \ldots \quad . \tag{2.3.1}$$

Let (Ω, F, μ) be the probability space which describes all the underlying uncertainties. Let $\underline{x}(0)$, $\underline{\xi}(k)$, $k = 0, 1, \ldots$ be independent Gaussian vectors with statistical laws:

$$\underline{x}(0) \sim G\left(\underline{x}_0, \underline{\Sigma}_0\right)$$

$$\underline{\xi}(k) \sim G\left(\underline{0}, \underline{R}(k)\right) \quad ; \quad k = 0, 1, \ldots \quad . \tag{2.3.2}$$

From (2.3.1), since $\underline{A}(k)$ is linear transformation, $\underline{x}(k)$ is also a Gaussian vector, $k = 0, 1, \ldots$.[3] Let f be some statistic on (Ω, F, μ), and let $F(f)$ denote the σ-algebra by f. Suppose that f is independent of $\xi(i)$, $i = k, k+1, \ldots$; then for $i \geq k$, by (2.2.12), we have,

$$\hat{\underline{x}}(i+1|f) = \underline{A}(i)\hat{\underline{x}}(i|f) \quad \text{a.s.} \quad ; \quad \hat{\underline{x}}(i|f) \triangleq E\{\underline{x}(i)|F(f)\} \tag{2.3.3}$$

and

$$\hat{\underline{\Sigma}}(i+1|f) = \underline{A}(i)\hat{\underline{\Sigma}}(i|f)\underline{A}'(i) + \underline{R}(i) \quad \text{a.s.} \quad ;$$
$$\hat{\underline{\Sigma}}(i|f) = E\{\underline{x}(i)\underline{x}'(i)|F(f)\} \quad . \tag{2.3.4}$$

Using (2.3.4) and (2.3.3), the conditional covariance of $\underline{x}(i)$ relative to $F(f)$, denoted by $\underline{\Sigma}^c(i|f)$, will satisfy

$$\underline{\Sigma}^c(i+1|f) = \underline{A}(i)\underline{\Sigma}^c(i|f)\underline{A}'(i) + \underline{R}(i) \qquad \text{a.s.} \quad . \tag{2.3.5}$$

In addition, if the conditional distribution of $\underline{x}(k)$ relative to $F(f)$ is Gaussian, then for all $i \geq k$, $\underline{x}(i)$ is a conditional Gaussian vector relative to $F(f)$. The statistical information of the statistic f is contained in $F(f)$, but the necessary statistical information of f _about_ the uncertainty of the future state of the system, $\underline{x}(i)$, $i \geq k$, is contained in the conditional distribution of $\underline{x}(k)$ relative to $F(f)$; and if it is Gaussian, $\hat{\underline{x}}(k|f)$ and $\hat{\underline{\Sigma}}(k|f)$ completely specify the conditional distribution of $\underline{x}(i)$, $i \geq k$, relative to $F(f)$. This is also referred to as the _Markov property_.

In the above discussion, the observation statistic is completely general. If the observation statistic is linear in $\underline{x}(j)$, $j = 0, 1, \ldots, k$, and some other Gaussian vectors, e.g., $f = \{\underline{y}(0), \ldots, \underline{y}(k)\}$, and

$$\underline{y}(i) = \underline{C}(i)\ x(i) + \underline{n}(i) \qquad i = 0, 1, \ldots, k \tag{2.3.6}$$

where $\underline{n}(i)$ is F-measurable, independent of $\underline{\xi}(j)$, $j \geq k$ and of Gaussian statistical law, $i = 0, 1, \ldots, k$; then $\underline{x}(k)$ is conditional Gaussian relative to $F(f)$.[1] There may be other kinds of statistical observations which will induce a conditional Gaussian law on $\underline{x}(k)$, but in this thesis, we shall only consider observation statistics of the type given by (2.3.6).

Consider a continuous linear stochastic system described by

$$\dot{\underline{x}}(t) = \underline{A}(t)\ \underline{x}(t) + \underline{\xi}(t) \tag{2.3.7}$$

where $\underline{A}(\cdot)$ is measurable in t, and is locally bounded.

Let (Ω, F, μ) be the basic probability space where $\underline{x}(t_0)$, $\underline{\xi}(t)$, $t \in [t_0, T]$ are statistics defined on Ω. The solution of (2.3.7), $\underline{x}(t)$, is defined as a process which satisfies the integral equation

$$\underline{x}(t) = \underline{x}(t_0) + \int_{t_0}^{t} \underline{A}(\tau)\underline{x}(\tau)d\tau + \int_{t_0}^{t} \underline{\xi}(\tau)d\tau \quad ; \quad t \in [t_0, T] \tag{2.3.8}$$

Let $\underline{\xi}(t)$ be a Gaussian White Noise Process with

$$E\left\{\int_{t_1}^{t_2} \underline{\xi}(\tau)d\tau\right\} = \underline{0} \qquad ; \qquad t_0 \leq t_1 < t_{0_L} \leq T \tag{2.3.9}$$

$$E\left\{\left(\int_{t_1}^{t_2} \underline{\xi}(\tau)d\tau\right)\left(\int_{t_1}^{t_2} \underline{\xi}(\tau)d\tau\right)'\right\} = \int_{t_1}^{t_2} \underline{R}(\tau)d\tau \qquad ; \qquad t_0 \leq t_1 < t_2 \leq T \ . \tag{2.3.10}$$

From (2.3.8), we see that for $T \geq t_2 > t_1 \geq t_0$

$$\underline{x}(t_2) = \underline{x}(t_1) + \int_{t_1}^{t_2} \underline{A}(\tau)\underline{x}(\tau)d\tau + \int_{t_1}^{t_2} \underline{\xi}(\tau)d\tau \ . \tag{2.3.11}$$

We shall always assume that $\underline{x}(t_0)$ and $\underline{\xi}(\tau)$, $\tau \varepsilon [t_0, t]$, are independent for all $t > t_0$. Therefore $\underline{x}(t_1)$ is independent of $\underline{\xi}(\tau)$, $\tau \varepsilon [t_1, t]$, for all $t > t_1$. We can find the solution of (2.3.11), $\underline{x}(\tau)$, $\tau \varepsilon [t_1, T]$, by successive approximation: $t \varepsilon [t_1, T]$

$$\underline{x}_0(t) = \underline{x}(t_1) \quad ; \quad \underline{x}_1(t) = \underline{x}(t_1) + \int_{t_1}^{t} \underline{A}(\tau)\underline{x}_0(\tau)d\tau + \int_{t_1}^{t} \underline{\xi}(\tau)d\tau \qquad (2.3.12)$$

and

$$\underline{x}_n(t) = \underline{x}(t_1) + \int_{t_1}^{t} \underline{A}(\tau)\underline{x}_{n-1}(\tau)d\tau + \int_{t_1}^{t} \underline{\xi}(\tau)d\tau \qquad n = 1, 2, \ldots \quad . \qquad (2.3.13)$$

By the assumptions on $\underline{A}(\cdot)$ and $\underline{R}(\cdot)$, this procedure will converge with $\underline{x}_n(t) \to \underline{x}(t)$, $t \varepsilon [t_1, T]$ a.s., and $\underline{x}(t)$ satisfies (2.3.11).

Let f be an observation statistic such that F(f) and $F(\underline{\xi}(\tau), \tau \varepsilon [t_1, T])$ are independent. Suppose that the conditional distribution of $\underline{x}(t_1)$ relative to f is Gaussian. Then from (2.3.12) and (2.3.13), we see that the conditional distribution of $\underline{x}_n(t)$ for a fixed $t \varepsilon [t_1, T]$ is Gaussian relative to f, $n = 0, 1, \ldots$, thus $\underline{x}(t)$, for a fixed $t \varepsilon [t_1, T]$, which satisfies (2.3.11) is also conditionally Gaussian relative to F(f). Therefore the complete statistical law of $\underline{x}(t)$ relative to f is described by its conditional mean and covariance.

From (2.3.13), we see that for all $n = 1, 2, \ldots$

$$E\{\underline{x}_n(t)\,|F(f)\} = \left\{\underline{1}+\int_{t_1}^{t} \underline{A}(\tau_1)\int_{t_1}^{\tau_1} \underline{A}(\tau_2)\ldots\int_{t_1}^{\tau_{n-1}} \underline{A}(\tau_n)d\tau_n\ldots d\tau_1\right\}\cdot E\{\underline{x}(t_1)\,|F(f)\} \text{ a.s.} \qquad (2.3.14)$$

$E\{\underline{x}_n(t)\underline{x}_n'(t)\,F(f)\} =$

$$\left(\underline{I}+\int_{t_1}^{t}\underline{A}(\tau_1)\int_{t_1}^{\tau_1}\underline{A}(\tau_2)\cdots\int_{t_1}^{\tau_{n-1}}\underline{A}(\tau_n)d\tau_n\cdots d\tau_1\right)\cdot E\{\underline{x}(t_1)\underline{x}'(t_1)\,F(f)\}\cdot$$

$$\left(\underline{I}+\int_{t_1}^{t}\underline{A}(\tau_1)\int_{t_1}^{\tau_1}\underline{A}(\tau_2)\cdots\int_{t_1}^{\tau_{n-1}}\underline{A}(\tau_n)d\tau_n\cdots d\tau_1\right)'+\int_{t_1}^{t}\underline{R}(\tau)d\tau$$

$$+\int_{t_1}^{t}\underline{A}(\sigma_1)\int_{t_1}^{\sigma_1}\underline{A}(\sigma_2)\cdots\int_{t_1}^{\sigma_{n-2}}\underline{A}(\sigma_{n-1})\int_{t_1}^{\sigma_{n-1}}\underline{R}(\sigma_n)d\sigma_n\cdots d\sigma_1$$

$$+\left(\int_{t_1}^{t}\underline{A}(\sigma_1)\int_{t_1}^{\sigma_1}\underline{A}(\sigma_2)\cdots\int_{t_1}^{\sigma_{n-2}}\underline{A}(\sigma_{n-1})\int_{t_1}^{\sigma_{n-1}}\underline{R}(\sigma_n)d\sigma_n\cdots d\sigma_1\right)'\quad \text{a.s.} \quad . \quad (2.3.15)$$

Since $\underline{x}_n(t) \to \underline{x}(t)$ a.s., $E\{\underline{x}_n(t)|F(f)\} \to E\{\underline{x}(t)|F(f)\}$ a.s., and $E\{\underline{x}_n(t)\underline{x}_n'(t)|F(f)\} \to E\{\underline{x}(t)\underline{x}'(t)|F(f)\}$ a.s. Equations (2.3.14) and (2.3.15) imply that as $n \to \infty$, $E\{\underline{x}(t)|F(f)\}$ and $E\{\underline{x}(t)\underline{x}'(t)|F(f)\}$ satisfy (a.s.):

$$\dot{\hat{\underline{x}}}(t|F(f)) = \underline{A}(t)\hat{\underline{x}}(t|F(f)) \qquad t \geq t_1 \quad ; \quad \hat{\underline{x}}(t|F(f)) \triangleq E\{\underline{x}(t)|F(f)\} \quad (2.3.16)$$

$$\dot{\hat{\underline{\Sigma}}}(t|F(f)) = \underline{A}(t)\hat{\underline{\Sigma}}(t|F(f)) + \hat{\underline{\Sigma}}(t|F(f))\underline{A}'(t) + \underline{R}(t) \qquad t \geq t_1 \quad ;$$

$$\hat{\underline{\Sigma}}(t|F(f)) \triangleq E\{\underline{x}(t)\underline{x}'(t)|F(f)\} \quad . \quad (2.3.17)$$

The conditional covariance of $\underline{x}(t)$, $t \geq t_1$, denoted by $\underline{\Sigma}^c(t|F(f))$ will then satisfy

$$\dot{\underline{\Sigma}}^c(t|F(f)) = \underline{A}(t)\underline{\Sigma}^c(t|F(f)) + \underline{\Sigma}^c(t|F(f))\underline{A}'(t) + \underline{R}(t) \quad \text{a.s.} \quad , \quad t \geq t_1 \quad . \; (2.2.18)$$

In the above, the observation statistic is completely general. If the conditional distribution of $\underline{x}(t_1)$ relative to f is <u>not</u> Gaussian, then $\underline{x}(t)$, $t \geq t_1$, will not be Gaussian for any fixed t; however the conditional mean and covariance of $\underline{x}(t)$, $t \geq t_1$, are still given by (2.3.16) and (2.3.18).

In this thesis, we shall assume that the observation statistic is of the form

$$\underline{y}(t) = \underline{C}(t)\, \underline{x}(t) + \underline{n}(t) \quad , \quad t \in [t_0, T] \tag{2.3.19}$$

where $\underline{n}(t)$, $t \in [t_0, T]$ is Gaussian white noise with

$$E\left\{\int_{t_1}^{t_2} \underline{n}(\tau)d\tau\right\} = \underline{0} \tag{2.3.20}$$

$$E\left\{\left(\int_{t_1}^{t_2} \underline{n}(\tau)d\tau\right)\left(\int_{t_1}^{t_2} \underline{n}(\tau)d\tau\right)'\right\} = \int_{t_1}^{t_2} \underline{Q}(\tau)d\tau \tag{2.3.21}$$

and $\underline{n}(t)$, $t \in [t_0, T]$ is independent of $\underline{\xi}(t)$, $t \in [t_0, T]$, and $\underline{x}(t_0)$. With such observation statistic, we see that $F_{t_1} \triangleq F(\underline{y}(\tau), \tau \in [t_0, t_1])$ is independent of $\underline{\xi}(\tau)$, $\tau \in [t_1, T]$; furthermore $\hat{\underline{x}}(t_1|F_{t_1})$ is F_{t_1}-measurable and $\hat{\underline{x}}(t_1|F_{t_1})$ is conditionally Gaussian if $\underline{x}(t_0)$ is Gaussian.[4] If $t_0 < t_1 < t_2 \ldots < t_n$, we have

$$F_{t_1} \subsetneq F_{t_2} \ldots \subsetneq F_{t_n} \subsetneq F \quad .$$

In the more general nonlinear case, the system is described by

$$\dot{\underline{x}}(t) = \underline{f}(t, \underline{x}(t)) + \underline{\xi}(t) \tag{2.3.22}$$

where $\underline{\xi}(t)$, $t \in [t_0, T]$, is a Gaussian white noise with statistical law (2.3.9), (2.3.10), and $\underline{f}(t, \underline{x}(t))$ is $F(\underline{\xi}(\tau), \tau \in [t_0, t])$-measurable, the solution of (2.3.22) is defined as the process which satisfies

$$\underline{x}(t) = \underline{x}(t_0) + \int_{t_0}^{t} \underline{f}(\tau, \underline{x}(\tau))d\tau + \int_{t_0}^{t} \underline{\xi}(\tau)d\tau \quad \text{a.s.} \quad t \in [t_0, T] \,. \tag{2.3.23}$$

If $f(t, \cdot)$ satisfies the Lipschitz condition

$$\left\| \underline{f}\left(t, \underline{x}_1\right) - \underline{f}\left(t, \underline{x}_2\right) \right\| < \alpha \left\| \underline{x}_1 - \underline{x}_2 \right\| \quad ; \quad \underline{x}_1, \underline{x}_2 \in R^n \qquad (2.3.24)$$

where α is some constant; then the method of successive approximation by setting $\underline{x}(t) = \underline{x}(t_0)$ and

$$\underline{x}_n(t) = \underline{x}(t_0) + \int_{t_0}^{t} \underline{f}(\tau, \underline{x}_{n-1}(\tau))d\tau + \int_{t_0}^{t} \underline{\xi}(\tau)d\tau \qquad n = 1, 2, \ldots \quad (2.3.25)$$

will converge almost surely to $\underline{x}(t)$, as $n \to \infty$.[6] The interpretation we used here is Itô's; the reason for adapting this interpretation is due to the rich mathematical properties one can deduce and utilize by using this interpretation. Itô's theory in stochastic differential equation will not be considered in here, the detail can be found in [1], [7], [8].

Let $\underline{x}(t)$ be a process described by (2.3.22) or (2.3.23); $\underline{x}(t)$ is called a diffusion process.[5],[7] Let $C(\cdot, \cdot)$ be defined on $T \times R^n$ with real scalar value, such that $\underline{C}_x(t, x)$, $C_t(t, x)$, and $\underline{C}_{xx}(t, x)$ are defined and continuous. The differential generator of $\underline{x}$ with respect to C is defined by

$$\mathfrak{L}(C(t, \underline{x})) \triangleq \lim_{s \downarrow t} (s - t)^{-1} E\{C(t + \Delta t, \underline{x}(t + \Delta t)) - C(t, \underline{x}(t) | \underline{x}(t) = \underline{x}\} \quad . \qquad (2.3.26)$$

If $\underline{x}(t)$, $t \in [t_0, T]$, satisfies (2.3.22), then [7],[8]

$$\mathfrak{L}(C(t, \underline{x})) = \frac{1}{2} \operatorname{tr}\left\{\underline{R}^{1/2}(t)'\underline{C}_{xx}(t, \underline{x})\underline{R}^{1/2}(t)\right\} + \underline{f}(t, \underline{x})'\underline{C}_x(t, \underline{x}) \quad . \qquad (2.3.27)$$

If in addition,

$$|C| + |C_t| + |\underline{x}||\underline{C}_x| + |\underline{x}|^2|\underline{C}_{xx}| \le k\left(1 + |\underline{x}|^2\right) \quad ; \quad (t, \underline{x}) \in T \times R^n \qquad (2.3.28)$$

then

$$C(t,\underline{x}(t)) = C(t_1,\underline{x}(t_1)) + \int_{t_1}^{t} [\mathcal{L}(C(\tau,\underline{x}(\tau)) + C_\tau(\tau,\underline{x}(\tau))]d\tau$$

$$+ \int_{t_1}^{t} \underline{C}_x(\tau,\underline{x}(\tau))'\underline{\xi}(\tau)d\tau \qquad (2.3.29)$$

where now the last integral must be interpreted in the sense of Itô.[17] Let F_1 be the sub-σ-algebra which is independent of $F(\underline{\xi}(\tau);\ \tau \in [t_1, T])$. Since [4]

$$E\left\{\int_{t_1}^{t} \underline{C}_x(\tau,\ \underline{x}(\tau))'\ \underline{\xi}(\tau)d\tau \,\middle|\, F_1\right\} = 0 \qquad (2.3.30)$$

we have from (2.3.29) the Itô's integration formula:[17],[4]

$$E\left\{C(t,\underline{x}(t))\middle| F_1\right\} = E\left\{C(t_1,\underline{x}(t_1))\middle| F_1\right\} + E\left\{\int_{t_1}^{t} \left[\mathcal{L}(C(\tau,\underline{x}(\tau)) + C_\tau(\tau,\underline{x}(\tau)\right]d\tau \,\middle|\, F_1\right\} . \qquad (2.3.31)$$

2.4 Stochastic Optimization

In this section, the mathematical tools for stochastic optimization problems are stated, and the outline of the proofs will be given. These stochastic optimization techniques will be used in later chapters to solve different stochastic control problems.

Since we shall be considering linear systems with Gaussian disturbances, the process which we shall control will be Gaussian. Thus an adequate description of the process is the evolution of its mean and covariance. As a result, we shall deal with a set of deterministic equations which describes the "trajectory" of the mean and covariance. In many cases, we can transform

a linear stochastic control problem into a deterministic control problem where the dynamics of the deterministic system are described by a set of matrix and vector differential equations. After making such transformation, the technique of the <u>matrix minimum principle</u> can be used to obtain necessary conditions for optimality,[9],[10] in the following way.

<u>Discrete Time Control Problem</u>:

A set of matrix and vector difference equations is given:

$$\left.\begin{aligned}\underline{X}(k+1) - \underline{X}(k) &= \underline{F}(k,\underline{X}(k),\underline{x}(k),\underline{U}(k)) \\ \underline{x}(k+1) - \underline{x}(k) &= \underline{f}(k,\underline{x}(k),\underline{X}(k),\underline{U}(k))\end{aligned}\right\} ; \quad \begin{aligned}&k = 0, 1, \ldots, N-1 \\ &\underline{X}(\cdot) = \underline{X}_0;\ \underline{x}(\cdot) = \underline{x}_0\end{aligned} \tag{2.4.1}$$

with $\underline{U}(k) \in S$, constrained control set, $\underline{X}(k) \in M_{nm}$, $\underline{x}(k) \in R^p$. Consider the scalar cost:

$$J = K(\underline{X}(N),\underline{x}(N)) + \sum_{k=0}^{N-1} L(k,\underline{U}(k),\underline{X}(k),\underline{x}(k)) \quad . \tag{2.4.2}$$

It is assumed that $\underline{F}(k,\cdot)$, $\underline{f}(k,\cdot)$, $K(\cdot)$ and $L(k,\cdot)$ satisfy the conditions required by the discrete minimum principle.[33] The control problem is to choose $\underline{U}^*(k)$, $K = 0, \ldots, N-1$, such that the cost (2.4.2) is minimized subject to the constraint (2.4.1) and $\underline{U}^*(k) \in S$, $k = 0, \ldots, N-1$.

Define the Hamiltonian function

$$\begin{aligned}H(\underline{X}(k),\underline{x}(k),\underline{P}(k+1),\underline{p}(k+1),\underline{U}(k)) &\triangleq L(k,\underline{U}(k),\underline{X}(k),\underline{x}(k)) \\ &+ \underline{f}'(k,\underline{x}(k),\underline{X}(k),\underline{U}(k))\underline{p}(k+1)+\mathrm{tr}\{\underline{F}(k,X(k),\underline{x}(k),\underline{U}(k))\underline{P}'(k+1)\}\end{aligned} \tag{2.4.3}$$

where $\underline{P}(k)$, $\underline{p}(k)$ are the costate associated with $\underline{X}(k)$ and $\underline{x}(k)$ respectively.

Theorem 2.4.1: (Matrix Minimum Principle: Discrete Time)

Let $\underline{U}^*(k)$, $k = 0, \ldots, N-1$ be the optimal control and $\underline{X}^*(k)$, $\underline{x}^*(k)$, $k = 0, \ldots, N$ be the optimal state; then there exists a costate matrix $\underline{P}^*(k)$ associated with $\underline{X}^*(k)$, and a costate vector $\underline{p}^*(k)$ associated with $\underline{x}^*(k)$ such that the following relations hold:

1) Canonical Equations:

$$\underline{X}^*(k+1) - \underline{X}^*(k) = \left.\frac{\partial H}{\partial \underline{P}(k+1)}\right|_* \quad ; \quad \underline{x}^*(k+1) - \underline{x}^*(k) = \left.\frac{\partial H}{\partial \underline{p}(k+1)}\right|_*$$

$$\underline{P}^*(k+1) - \underline{P}^*(k) = -\left.\frac{\partial H}{\partial \underline{X}(k)}\right|_* \quad ; \quad p^*(k+1) - \underline{p}^*(k) = -\left.\frac{\partial H}{\partial \underline{x}(k)}\right|_* \tag{2.4.4}$$

2) Boundary Conditions:

$$\underline{X}^*(0) = \underline{X}_0 \quad ; \quad \underline{x}^*(0) = \underline{x}_0 \tag{2.4.5}$$

$$\underline{P}^*(N) = \frac{\partial K(\underline{X}^*(N),\underline{x}^*(N))}{\partial \underline{X}^*(N)} \quad ; \quad \underline{p}^*(N) = \frac{\partial K(\underline{X}^*(N),\underline{x}^*(N))}{\partial \underline{x}^*(N)} \tag{2.4.6}$$

3) Minimization of the Hamiltonian:

For every $\underline{U} \in S$, and for each $k = 0, 1, \ldots, N-1$

$$H\Big(\underline{X}^*(k),\underline{x}^*(k),\underline{P}^*(k+1),\underline{p}^*(k+1),\underline{U}^*(k)\Big) \le H\Big(\underline{X}^*(k),\underline{x}^*(k),\underline{P}^*(k+1),\underline{p}^*(k+1),\underline{U}\Big) \tag{2.4.7}$$

Continuous Time Control Problems:

A set of matrix and vector differential equations is given:

$$\dot{\underline{X}}(t) = \underline{F}(t,\underline{X}(t),\underline{x}(t),\underline{U}(t)) \quad ; \quad \underline{X}(t_0) = \underline{X}_0$$

$$\dot{\underline{x}}(t) = \underline{f}(t,\underline{x}(t),\underline{X}(t),\underline{U}(t)) \quad ; \quad \underline{x}(t_0) = \underline{x}_0 \tag{2.4.8}$$

with $\underline{U}(t) \in S$, constrained control set, $\underline{X}(t) \in M_{nm}$, $\underline{x}(t) \in R^p$. Consider the scalar cost:

$$J = K(\underline{X}(T),\underline{x}(T)) + \int_{t_0}^{T} L(t,\underline{X}(t),\underline{x}(t),\underline{U}(t))dt \quad ; \quad T \text{ fixed} \quad . \quad (2.4.9)$$

The usual differentiability conditions for $\underline{F}(\cdot)$, $\underline{f}(\cdot)$, $K(\cdot)$, and $L(\cdot)$ are assumed to be satisfied. The control problem is to choose $\underline{U}^*(t)$, $t \in [t_0, T]$, such that the cost (2.4.9) is minimized subject to the constraint (2.4.8) and $\underline{U}^*(t) \in S$.

Define the Hamiltonian function

$$H(\underline{X}(t),\underline{x}(t),\underline{P}(t),\underline{p}(t),\underline{U}(t)) \triangleq L(t,\underline{X}(t),\underline{x}(t),\underline{U}(t)) + \underline{f}'(t,\underline{x}(t),\underline{X}(t),\underline{U}(t))\underline{p}(t) + \mathrm{tr}\{\underline{F}(t,\underline{X}(t),\underline{x}(t),\underline{U}(t))\underline{P}'(t)\} \quad (2.4.10)$$

where $\underline{P}(t)$, $\underline{p}(t)$ are the costate associated with $\underline{X}(t)$ and $\underline{x}(t)$ respectively.

<u>Theorem 2.4.2</u>: (Matrix Minimum Principle: Continuous Time)

Let $\underline{U}^*(t)$, $t \in [t_0, T]$, be the optimal control and $\underline{X}^*(t)$, $\underline{x}^*(t)$, $t \in [t_0, T]$, be the optimal state, then there exist costates $\underline{P}^*(t)$, $\underline{p}^*(t)$ such that the following conditions hold:

1) Canonical Equations:

$$\dot{\underline{X}}^*(t) = \left.\frac{\partial H}{\partial \underline{P}(t)}\right|_* \quad ; \quad \dot{\underline{x}}^*(t) = \left.\frac{\partial H}{\partial \underline{p}(t)}\right|_* \quad (2.4.11)$$

$$\dot{\underline{P}}^*(t) = -\left.\frac{\partial H}{\partial \underline{X}(t)}\right|_* \quad ; \quad \dot{\underline{p}}^*(t) = -\left.\frac{\partial H}{\partial \underline{x}(t)}\right|_* \quad (2.4.12)$$

2) Boundary Conditions:

$$\underline{X}^*(t_0) = \underline{X}_0 \quad ; \quad \underline{x}^*(t_0) = \underline{x}_0 \tag{2.4.13}$$

$$\underline{P}^*(T) = \frac{\partial K(\underline{X}^*(T), \underline{x}^*(T))}{\partial \underline{X}^*(T)} \quad ; \quad \underline{p}^*(T) = \frac{\partial K(\underline{X}^*(T), \underline{x}^*(T))}{\partial \underline{x}^*(T)} \tag{2.4.14}$$

3) Minimization of the Hamiltonian:

$$H\left(\underline{X}^*(t), \underline{x}^*(t), \underline{P}^*(t), \underline{p}^*(t), \underline{U}^*(t)\right) \le H\left(\underline{X}^*(t), \underline{x}^*(t), \underline{P}^*(t), \underline{p}^*(t), \underline{U}\right) \tag{2.4.15}$$

for all $\underline{U} \in S$ and $t \in [t_0, T]$.

The matrix minimum principle (both discrete and continuous) is a just straightforward extension of the vector minimum principle, Holtzman and Halkin [33], Pontryagin, et al. [11]. Theoretically, the justification of the matrix minimum principle hinges on the existence of a mapping from M_{nm} to R^{nm}. The details were carried out by Tse [9]; see also Athans [12].

The matrix minimum principle only provides us with necessary conditions for optimality. A control and its corresponding state trajectory which satisfies the matrix minimum principle will be called extremal control and extremal state trajectory. If one can prove the existence of optimal controls and the uniqueness of extremal controls, the matrix minimum principle also served as a sufficient condition for optimality. But, in general, the matrix minimum principle does not provide sufficiency. It will be convenient if one can have some sufficient conditions for optimality, so that one can easily test to see whether an extremal control is optimal or not. It turns out that to look for sufficient conditions, it is often easier (and more general) if we consider the original stochastic control problem without transforming it to deterministic description in terms of mean and covariance.

Discrete Time Stochastic Control Problems:

A discrete time stochastic process is described by

$$\underline{x}(k+1) = \underline{f}(k,\underline{x}(k),\underline{u}(k)) + \underline{\xi}(k) \qquad k = k_0, k_0+1, \ldots \qquad (2.4.16)$$

with $\underline{x}(k) \in R^n$, $\underline{u}(k) \in R^r$. Let $\underline{x}(0)$, $\underline{\xi}(k)$, $k = 0, 1, \ldots$ be independent Gaussian Vectors defined on the basic probability space with statistical law (2.3.2).

Let $U(k_0, k) \triangleq \{\underline{u}(k_0),\underline{u}(k_0+1),\ldots,\underline{u}(k)\}$ denotes the control sequence, and $g(k_0)$, $g(k_0+1,\underline{u}(k_0))$, $g(k_0+2,U(k_0,k_0+1))$, ..., $g(k,U(k_0, k-1))$, ... is a sequence of observation statistics which depends on control sequence, such that for all control sequences $F(k,U(k_0,k-1) \subset F(k+1,U(k_0,k))$, where $F(g(k,U(k_0,k-1)) \triangleq F(k,U(k_0,k-1))$. Let $\{\underline{x}_{U(k_0,k-1)}(k)\}_{k=k_0}^{N}$ be the process described by (2.4.16) when control sequence $U(k_0,N-1)$ is applied. Assume that $\underline{x}_{U(k_0,k-1)}(k)$ is $F(g(k,U(k_0,k-1))$-measurable when the control is restricted to be of the form:

$$\underline{u}(k) = \underline{\phi}(k,g(k,U(k_0,k-1)) \in S \qquad \text{(a.s.)} \quad . \qquad (2.4.17)$$

The control problem is to find air optimal control law $\underline{\phi}^*(k, U^*(k,_0,k-1))$ such that

$$J(U(k_0,N-1)|F(g(k_0)) = E\left\{K(\underline{x}(N)) + \sum_{k=k_0}^{N-1} L(k,\underline{x}(k),\underline{u}(k))|F(g(k_0))\right\} \quad (2.4.18)$$

is minimized subject to (2.4.16).

Theorem 2.4.3: (Optimality Criteria: Discrete Time)

Suppose that there exists a control strategy $\underline{\phi}^*(i,\cdot)$: $I_{[k_0,N-1]} x R^n \to S$ and a scalar function $C(\cdot,\cdot)$: $I_{[k_0,N-1]} x R^n \to R^1$ such that almost surely,

1) $C(N,\underline{x}) = K(\underline{x})$ (2.4.19)

2) $$0 = E\{L(k,\underline{x},\underline{\phi}^*(k,\underline{x}))+C(k+1,\underline{x}^*(k+1,\underline{x})|F(k,U(k_0,k-1))\}-C(k,\underline{x}) \tag{2.4.20}$$

$$\le E\{L(k,\underline{x},\underline{u})+C(k+1,\underline{x}_u(k+1,\underline{x})|F(k,U(k_0,k-1)\}-C(k,\underline{x}) \quad \text{a.s.}$$

$$k = k_0,\ k_0+1,\ \ldots$$

where for $k = k_0, k_0 + 1, \ldots, N - 1$

$$\underline{x}^*(k+1,\underline{x}) = \underline{A}(k)\underline{x} + \underline{B}(k)\underline{\phi}^*(k,\underline{x}) + \underline{\xi}(k) \tag{2.4.21}$$

$$\underline{x}_u(k+1,\underline{x}) = \underline{A}(k)\underline{x} + \underline{B}(k)\underline{u} + \underline{\xi}(k) \quad ; \quad \underline{u} = \underline{\phi}(k,g_k) \in S$$

i.e., $\underline{u}$ is any permissible control value at k. Then the control law $\underline{\phi}^*(\cdot,\cdot)$ is optimal and $C(k_0,\underline{x}(k_0))$ is the optimal cost a.s.

Proof: Let $\underline{x}^*(k), k = k_0,\ldots,N$ be the random vectors which satisfy the difference equation (a.s.)

$$\underline{x}^*(k+1) = \underline{A}(k)\underline{x}^*(k) + \underline{B}(k)\underline{\phi}^*(k,\underline{x}^*(k)) + \underline{\xi}(k) \quad ;$$
$$\underline{x}^*(k_0) = \underline{x}(k_0) \quad . \tag{2.4.23}$$

By using lemma 2.2.6, we have from (2.4.20) that: (a.s.)

$$C(k_0,\underline{x}(k_0)) = E\{L(k_0,\underline{x}(k_0),\underline{\phi}^*(k_0,\underline{x}(k_0))+C(k_0+1,\underline{x}^*(k+1)|F(g(k_0))\}$$

$$= E\{L(k_0,\underline{x}(k_0))\underline{\phi}^*(k_0,\underline{x}(k_0))|F(g(k_0))\}$$

$$+E\{E\{L(k_0+1,\underline{x}^*(k_0+1))\underline{\phi}^*(k_0+1,\underline{x}^*(k_0+1))$$

$$+C(k_0+2,\underline{x}^*(k_0+2))|F(k_0+1,\underline{\phi}^*(k_0,g(k_0))\}|F(g(k_0))\}$$

$$= E\sum_{k=k_0}^{k_0+1} L(k,\underline{x}^*(k),\underline{\phi}^*(k,\underline{x}^*(k))|F(g(k_0))$$

$$+E\{C(k_0+2,\underline{x}^*(k_0+2))|F(g(k_0))\} \quad . \tag{2.4.24}$$

Inductively and keeping (2.4.19) in mind, we have

$$C(k_0,\underline{x}(k_0)) = E\left\{K(\underline{x}^*(N))+\sum_{k=k_0}^{N-1} L(k,\underline{x}^*(k),\underline{\phi}^*(k,\underline{x}^*(k)))\,|\,F(g(k_0))\right\} \quad \text{a.s.} \qquad (2.4.25)$$

Now let $U^o = \{\underline{u}^o(i)\}_{i=0}^{N-1}$ be any admissible control law of the form (2.4.17). Let $\underline{x}^o(k)$, $k = k_0, k_0 + 1, \ldots, N$ be the random vectors which satisfy (a.s.)

$$\underline{x}^o(k + 1) = \underline{A}(k)\underline{x}^o(k) + \underline{B}(k)\underline{u}^o(k) + \underline{\xi}(k) \quad ; \quad \underline{x}^o(k_0) = \underline{x}(k_0) \quad . \qquad (2.4.26)$$

By (2.4.17), $\underline{u}^o(k)$ is $F(k,U^o(k_0,k - 1))$-measurable. Using lemma 2.2.6 and the inequality of (2.4.20), we have (a.s.),

$$\begin{aligned} C(k_0,\underline{x}(k_0)) &\le E\{L(k_0,\underline{x}(k_0),\underline{u}^o(k)+C(k_0+1,\underline{x}^o(k_0+1))\,|\,F(g(k_0))\} \\ &\le E\{L(k_0,\underline{x}(k_0)),\underline{u}^o(k)\,|\,F(g(k_0))\}+E\{E\{L(k_0+1,\underline{x}^o(k_0+1),\underline{u}^o(k_0+1)) \\ &\quad +C(k_0+2,\underline{x}^o(k_0+2))\,|\,F(k_0+1,\underline{u}^o(k_0))\}\,|\,F(g(k_0))\} \\ &= E\left\{\sum_{k=k_0}^{k_0+1} L(k,\underline{x}^o(k),\underline{u}^o(k))\,|\,F(g(k_0))\right\}+E\{C(k_0+2,\underline{x}^o(k_0+2))\,|\,F(g(k_0))\} \quad . \end{aligned} \qquad (2.4.27)$$

Inductively and using (2.4.19), we have

$$C(k_0,\underline{x}(k_0)) \le E\left\{K(\underline{x}^o(N)) + \sum_{k=k_0}^{N-1} L(k,\underline{x}^o(k),\underline{u}^o(k))\,|\,F(g(k_0))\right\} \quad \text{a.s.} \qquad (2.4.28)$$

Combining (2.4.25) and (2.4.28) we have the assertion of the theorem.

Continuous Time Stochastic Control Problem:

A continuous time process is described by

$$\dot{\underline{x}}(t) = \underline{f}(t,\underline{x}(t)) + \underline{B}(t)\underline{u}(t) + \underline{\xi}(t) \qquad (2.4.29)$$

where $\underline{x}(t) \in R^n$, $\underline{f}(t,\ )$ satisfies the Lipschitz condition (2.3.18). $\underline{x}(t_0) \sim G(\underline{x}_0, \Sigma_0)$ and $\underline{\xi}(\tau)$, $\tau \in [t_0,T]$ is white Gaussian noise with statistical law (2.3.9) and (2.3.10). Denote the control $U[t_0,t) = \{\underline{u}(\tau), \tau \in [t_0,t)\}$. Let $g(t,U[t_0,t))$ be an observation statistic such that if $t_1 < t_2$, $F(t_1,U[t_0,t)) \doteq F(g(t_0 U[t_0,t_1)) \subset F(g(t_2,U[t_0,t_2)) \doteq F(t_2 U[t_0,t_2))$; at $t = t_0$, $g(t_0,U[t_0,t_0)) \doteq g(t_0)$ and is independent of control. Let $\{\underline{x}_{U[t_0,t]}(t),\ t \in [t_0,T]\}$ be the process described by (2.4.29) when $U[t_0,T)$ is applied. We assume that $\underline{x}_{U[t_0,t)}(t)$ is $F(t,U[t_0,t))$-measurable when the control is restricted to be of the form

$$\underline{u}(t) = \underline{\phi}(t,g(t,U[t_0,t))) \in S \quad \text{(a.s.)} \quad . \tag{2.4.30}$$

The control problem is to find optimal control law of the form (2.4.30) such that the cost

$$J(U[t_0,t])|F(g(t_0)) = E\left\{K(\underline{x}(T) + \int_{t_0}^{T} L(t,\underline{x}(t),\underline{u}(t))dt|F(g(t_0))\right\} \tag{2.4.31}$$

is minimized subject to (2.4.29).

For a fixed control $U^o[t_0,T)$ of the form (2.4.30), we have a fixed diffusion process described by

$$\dot{\underline{x}}(t) = \underline{f}(t,\underline{x}(t)) + \underline{B}(t)\underline{u}^o(t) + \underline{\xi}(t) \tag{2.4.32}$$

and we can associate with $U^o[t_0,T]$ a fixed differential generator $\mathcal{L}_{u^o}(\cdot)$. Let $C(t,\underline{x})$ be a scalar function, we have

$$\mathcal{L}_{u^o}(C(t,\underline{x})) = \frac{1}{2}\,\mathrm{tr}\{\underline{R}^{1/2}(t)\underline{C}_{xx}(t,\underline{x})\underline{R}^{1/2}(t)\}+(\underline{f}(t,\underline{x})+\underline{B}(t)\underline{u}^o(t))' \cdot \underline{C}_x(t,\underline{x}) \quad . \tag{2.4.33}$$

Theorem 2.4.2: (Optimality Criteria: Continuous Time)

Suppose there exists a control law $\underline{\phi}^*(\cdot,\cdot): [t_0,T) \times \underline{R}^n \to S$ with $\underline{\phi}^*(t,\cdot)$ satisfying the Lipschitz condition, and a function $C(\cdot,\cdot): [t_0,T) \times R^n \to R^1$ such that

1) $C(t,\underline{x})$, $C_t(t,\underline{x})$, $\underline{C}_x(t,\underline{x})$, $\underline{C}_{xx}(t,\underline{x})$ are continuous and for some k

$$|C| + |C_t| + |\underline{x}||\underline{C}_x| + |\underline{x}|^2|\underline{C}_{xx}| < k(1+|\underline{x}|^2) \qquad (t,\underline{x})\varepsilon[t_0,T)xR^n \quad (2.4.34)$$

2) $C(T,\underline{x}) = K(\underline{x})$ a.s. (2.4.35)

3) $$0 = C_t(t,\underline{x}) + E\{\mathcal{L}_{\phi^*}(C(t,\underline{x}) + L(t,\underline{x},\underline{\phi}^*(t,\underline{x}))|F(t,U[t_0,t))\}$$

$$\leq C_t(t,\underline{x}) + E\{\mathcal{L}_{\phi}(C(t,\underline{x}) + L(t,\underline{x},\underline{\phi}(t,\underline{x}))|F(t,U[t_0,t))\} \quad \text{a.s.} \quad (2.4.36)$$

for all $(t,\underline{x})\varepsilon[t_0,T]xR^n$, and $\underline{\phi}(t,\cdot)$ satisfies Lipschitz condition. Then $\underline{\phi}^*(t,\underline{x}(t))$ is the optimal control law and $C(t_0,\underline{x}(t_0))$ is the optimal cost a.s.

Proof: Let $\underline{x}^*(t)$ be the resulting diffusion process described by (2.4.32) where we adapt control law $\underline{\phi}^*(\cdot,\cdot)$. Using the Itô's integration formula (2.3.31) applying to the process $\underline{x}^*(t)$, we have

$$E\{C(t_0,\underline{x}^*(t_0))|F(g(t_0))\} = C(t_0,\underline{x}^*(t_0)) = E\{C(T,\underline{x}^*(T))|F(g(t_0))\}$$
$$- E\left\{\int_{t_0}^{T} [\mathcal{L}_{\phi^*}(C(\tau,\underline{x}^*(\tau)))\right.$$
$$\left. + C_\tau(\tau,\underline{x}^*(\tau))]d\tau|F(g(t_0))\right\} \quad \text{a.s.} \quad (2.4.37)$$

By lemma 2.2.6 and Equation (2.4.35), (2.4.36), we obtain

$$C(t_0,\underline{x}^*(t_0)) = E\{E\{C(T,\underline{x}^*(T))|F(\underline{x}^*(T))\}|F(g(t_0))\}$$

$$- E\Big\{E\Big\{\int_{t_0}^{T} [\mathcal{L}_{\rho^*}(C(\tau,\underline{x}^*(\tau)))$$

$$+ C_\tau(\tau,\underline{x}^*(\tau))]d\tau\, |F(t,U[t_0,t)) \Big\}|F(g(t_0))\Big\}$$

$$= E\left\{K(\underline{x}^*(T)) + \int_{t_0}^{T} L(\tau,\underline{x}^*(\tau),\underline{\rho}^*(\tau,\underline{x}^*(\tau))d\tau\, |F(g(t_0))\right\}$$

$$\text{a.s.} \qquad (2.4.38)$$

Let $\underline{\rho}(t,\cdot)$ be <u>any</u> control law which satisfies the Lipschitz condition, and $\underline{x}^o(t)$ be the resulting diffusion process described by (2.4.32) when control law $\underline{\rho}^o(t,\cdot)$ is used. The Itô's integration formula, applied now to the process $\underline{x}^o(t)$, gives us

$$C(t_0,\underline{x}^o(t_0)) = E\{C(T,\underline{x}^o(T))|F(g(t_0))\} - E\Big\{\int_{t_0}^{T} [\mathcal{L}_{\rho^o}(C(\tau,\underline{x}^o(\tau)))$$

$$+ C_\tau(\tau,\underline{x}^o(\tau))]d\tau\, |F(g(t_0))\Big\}$$

$$\leq E\Big\{K(\underline{x}^o(T)) + \int_{t_0}^{T} L(\tau,\underline{x}^o(\tau)\underline{\rho}^o(\tau,\underline{x}(\tau))d\tau\, |F(g(t_0))\Big\} \quad \text{a.s.}$$

$$(2.4.39)$$

where the last inequality comes from the inequality part of (2.4.36), also lemma 2.2.6 is being used in deriving (2.4.39). Note that $C(\cdot,\cdot)$ is defined on $[t_0,T) x R^n$; now equations (2.4.38) and (2.4.39) yield the statement of the theorem.

2.5 <u>Generalized Matrix Riccati Difference Equations</u>

For a given sequence of matrices $V \triangleq \{\underline{V}(k)\}_{k=k_0}^{\infty}$, $\underline{V}(k) \in M_{nm}$, $k = k_0$, $k_0 + 1, \ldots$, let $\{\underline{P}_V(k,k_0;\underline{F})\}_{k=k_0}^{\infty}$ be the sequence which satisfies the linear matrix difference equation:

$$\underline{P}_v(k+1,k_0;\underline{F}) = (\underline{A}(k)-\underline{V}(k)\underline{D}(k)\underline{A}(k))\underline{P}_v(k,k_0;\underline{F})(\underline{A}(k)-\underline{V}(k)\underline{D}(k)\underline{A}(k))'$$

$$+(\underline{I}_n-\underline{V}(k)\underline{D}(k))\underline{Q}_1(k)(\underline{I}_n-\underline{V}(k)\underline{D}(k))'+\underline{Q}_2(k)-\underline{Q}_1(k)+\underline{V}(k)\underline{R}(k)\underline{V}'(k)$$

$$\underline{P}_v(k_0,k_0;\underline{F}) = \underline{F} \tag{2.5.1}$$

where $\underline{A}(k) \in M_{nm}$, $\underline{D}(k) \in M_{mn}$ are bounded uniformly for all k. We assume that $\underline{F}$, $\underline{Q}_1(k)$, $\underline{Q}_2(k)$ are symmetric nonnegative definite nxn matrices with $\underline{Q}_1(k) \le \underline{Q}_2(k)$, $k = k_0, k_0+1, \ldots$, and that $\underline{R}(k)$ is symmetric nonnegative definite mxm matrix. Since (2.5.1) is linear, therefore for any arbitrary $\underline{V}(k) \in M_{nm}$, $k = k_0, k_0+1, \ldots$, $\{\underline{P}_v(k,k_0;\underline{F})\}_{k=k_0}^{\infty}$ exists, is unique and $\underline{P}_v(k,k_0;\underline{F}) \ge 0$, $k = k_0, k_0+1, \ldots$.

When $\underline{V}(k)$ ranges over M_{nm}, $k = k_0, k_0+1, \ldots$, we generate a solution set $\mathcal{B}_{k_0} = \{(\underline{P}_v(k,k_0;\underline{F}))_{k=k_0}^{\infty} \mid \underline{V}(k) \in M_{nm}, k = k_0, k_0+1, \ldots\}$. All elements in the solution set $\mathcal{B}_{k_0}$ is a sequence of symmetric nonnegative definite nxn matrices.

<u>Definition 2.5.1</u>: (Minimal sequence) Let $\mathcal{S}_{k_0} = \{(\underline{M}(k))_{k=k_0}^{\infty} \mid \underline{M}(k) \in M_{nn}$, $\underline{M}(k) \ge \underline{0}$, $k = k_0, k_0+1, \ldots\}$; an element $(\underline{M}^o(k))_{k=k_0}^{\infty} \in \mathcal{S}_{k_0}$ is called a minimal sequence with respect to $\mathcal{S}_{k_0}$ if for all $(\underline{M}(k))_{k=k_0}^{\infty} \in \mathcal{S}_{k_0}$, we have $\underline{M}^o(k) \le \underline{M}(k)$, $k = k_0, k_0+1, \ldots$.

For a given set $\mathcal{S}_{k_0}$, a minimal sequence may not exist; but if it exists, it must be unique. In the following, it will be shown that the solution set $\mathcal{B}_{k_0}$ has a unique minimal sequence.

Let us define the matrix $\underline{\Psi}_k(\underline{V},\underline{P})$ by $(k = k_0, k_0+1, \ldots)$

$$\underline{\Psi}_k(\underline{V},\underline{P}) = (\underline{A}(k)-\underline{V}\,\underline{D}(k)\underline{A}(k))\underline{P}(\underline{A}(k)-\underline{V}\,\underline{D}(k)\underline{A}(k))'+(\underline{I}_n-\underline{V}\,\underline{D}(k))\underline{Q}_1(k)(\underline{I}_n-\underline{V}\,\underline{D}(k))'$$

$$+\underline{Q}_2(k)-\underline{Q}_1(k)+\underline{V}\,\underline{R}(k)\underline{V}' \tag{2.5.2}$$

where $\underline{V} \in M_{nm}$, $\underline{P} \in M_{nn}$. For fixed $\underline{P} \in M_{nn}$, define the set $U_k(\underline{P}) = \{\underline{V} \in M_{nm} : (*)_k$ is satisfied$\}$ where the condition $(*)_k$ is given by

$$(*)_k: \quad \underline{\tilde{V}}(\underline{R}(k)+\underline{D}(k)(\underline{Q}_1(k)+\underline{A}(k)\underline{P}\ \underline{A}'(k))\underline{D}'(k)) = (\underline{A}(k)\underline{P}\ \underline{A}'(k)-\underline{Q}_1(k))\underline{D}'(k) \quad . \quad (2.5.3)$$

We have the following lemma:

<u>Lemma 2.5.2</u>: (Minimum property) Let $\underline{P} \in M_{nn}$, and $\underline{P} \geq \underline{0}$; if $\underline{\tilde{V}} \in U_k(\underline{P})$, then for all $\underline{V} \in M_{nm}$:

$$\underline{\Phi}_k(\underline{V},\underline{P}) \geq \underline{\Phi}_k(\underline{\tilde{V}},\underline{P}) \quad ; \quad k = k_0,\ k_0+1, \ldots \quad . \quad (2.5.4)$$

<u>Proof</u>: Let us denote

$$\underline{\hat{R}}(k,\underline{P}) = \underline{R}(k) + \underline{D}(k)(\underline{Q}_1(k) + \underline{A}(k)\underline{P}\ \underline{A}'(k))\underline{D}'(k) \qquad (2.5.5)$$
$$k = k_0,\ k_0+1, \ldots \quad .$$

The condition $(*)_k$ can now be written as

$$(*)'_k \qquad \underline{\tilde{V}}\ \underline{\hat{R}}(k,\underline{P}) = (\underline{A}(k)\underline{P}\ \underline{A}'(k) - \underline{Q}_1(k))\ \underline{D}'(k)$$
$$k = k_0,\ k_0+1, \ldots \quad .$$

Let $\underline{\tilde{V}} \in U_k(\underline{P})$, $\underline{\tilde{V}}$ must satisfy $(*)'_k$; and so for $\underline{V} \in M_{nm}$, we have

$$\underline{\Phi}_k(\underline{\tilde{V}},\underline{P})+(\underline{V}-\underline{\tilde{V}})\underline{\hat{R}}(k,\underline{P})(\underline{V}-\underline{\tilde{V}})'$$

$$= \underline{A}(k)\underline{P}\ \underline{A}'(k)+\underline{Q}_1(k)-\underline{\tilde{V}}\ \underline{D}(k)[\underline{A}(k)\underline{P}\ \underline{A}'(k)+\underline{Q}_1(k)]-[\underline{A}(k)\underline{P}\ \underline{A}'(k)+\underline{Q}_1(k)]\underline{D}'(k)\underline{V}'(k)$$

$$+2\underline{\tilde{V}}\ \underline{\hat{R}}(k,\underline{P})\underline{\tilde{V}}'+\underline{Q}_2(k)-\underline{Q}_1(k)+\underline{V}\ \underline{\hat{R}}(k,\underline{P})\underline{V}'-\underline{\tilde{V}}\ \underline{\hat{R}}(k,\underline{P})\underline{V}'-\underline{V}\ \underline{\hat{R}}(k,\underline{P})\underline{\tilde{V}}'$$

$$= \underline{A}(k)\underline{P}\ \underline{A}'(k)+\underline{Q}_1(k)+\underline{Q}_2(k)+\underline{V}\ \underline{\hat{R}}(k,\underline{P})\underline{V}'-(\underline{A}(k)\underline{P}(k)\underline{A}'(k)+\underline{Q}_1(k))\underline{D}'(k)\underline{V}'$$

$$-\underline{V}\ \underline{D}(k)(\underline{Q}_1(k)+\underline{A}(k)\underline{P}\ \underline{A}'(k)) = \underline{\Phi}_k(\underline{V},\underline{P}) \qquad k = k_0,\ k_0+1, \ldots \quad . \quad (2.5.6)$$

Since $\underline{R}(k,\underline{P}) \geq \underline{0}$, (2.5.4) follows from (2.5.6).

An immediate consequence of the lemma is that

$$\underline{\Phi}_k(\underline{V}_1,\underline{P}) = \underline{\Phi}_k(\underline{V}_2,\underline{P}) \quad , \quad \underline{V}_1,\underline{V}_2 \in \mathcal{V}_k(\underline{P}) \quad . \tag{2.5.7}$$

Theorem 2.5.3: There exists a unique minimal sequence, $\{\underline{P}^o(k,k_0;\underline{F})\}_{k=k_0}^{\infty}$ with respect to the solution set $\mathcal{R}_{k_0}$.

Proof: Let us construct the sequence $\{\underline{P}^o(k,k_0;\underline{F})\}_{k=k_0}^{\infty}$ as follows: Set $\underline{P}^o(k_0,k_0;\underline{F}) = \underline{F}$, and choose $\underline{V}^o(k_0) \in \mathcal{V}_{k_0}(\underline{F})$. Such a $\underline{V}^o(k_0)$ may not be unique, but by (2.5.7) and (2.5.1), this gives rise to a unique $\underline{P}^o(k_0 + 1,k_0;\underline{F}) = \underline{\Phi}_{k_0}(\underline{V}^o(k_0),\underline{F})$, $\underline{V}^o(k_0) \in \mathcal{V}_{k_0}(\underline{F})$.

Assume that $\underline{V}^o(k)$, k_0, $k_0 + 1$, ..., $k_0 + i$, have been chosen inductively with $\underline{V}^o(k) \in \mathcal{V}_k(\underline{P}^o(k,k_0;\underline{F})$ and a unique sequence

$$\underline{P}^o(k+1,k_0;\underline{F}) = \underline{\Phi}_k(\underline{V}^o(k),\underline{P}^o(k,k_0;\underline{F})) \quad , \quad k = k_0,k_0+1,\ldots,k_0+i \tag{2.5.8}$$

has been constructed. Choose $\underline{V}^o(k_0 + i + 1) \in \mathcal{V}_{k_0+i+1}(\underline{P}^o(k_0 + i + 1,k_0;\underline{F}))$. By (2.5.7) and (2.5.1), this gives a unique

$$\underline{P}^o(k_0+i+2,k_0;\underline{F}) = \underline{\Phi}_{k_0+i+1}(\underline{V}^o(k_0 + i + 1),\underline{P}^o(k_0+i+1,k_0;\underline{F})) \quad . \tag{2.5.9}$$

The sequence $\{\underline{P}^o(k,k_0;\underline{F})\}_{k=k_0}^{\infty}$ thus constructed is unique.

Let $\underline{V} = \{\underline{V}(k)\}_{k=k_0}^{\infty}$ be an arbitrary sequence with $\underline{V}(k) \in M_{nm}$, $k = k_0$, $k_0 + 1$, ... By lemma 2.5.2,

$$\underline{P}_V(k_0+1,k_0;\underline{F}) \geq \underline{P}^o(k_0+1,k_0;\underline{F}) \quad . \tag{2.5.10}$$

Assume that for $i \geq 1$,

$$\underline{P}_V(k_0+i,k_0;\underline{F}) \geq \underline{P}^o(k_0+i;k_0;\underline{F}) \quad . \tag{2.5.11}$$

From (2.5.1) and (2.5.2), w[illegible]r a given $\underline{V} \in M_{nm}$:

$$\Phi_k(\underline{V},\underline{P}_1) \geq \Phi_k(\underline{V},\underline{P}_2) \quad \text{if} \quad \underline{P}_1 \geq \underline{P}_2 \geq \underline{0} \quad , \quad k = k_0,k_0+1,\dots \quad . \tag{2.5.12}$$

Combining (2.5.1), (2.5.2), (2.5.4), (2.5.11), and (2.5.12), we have

$$\begin{aligned}\underline{P}_v(k_0+i+1,k_0;\underline{F}) &= \Phi_{k_0+i}(\underline{V}(k_0+i),\underline{P}_v(k_0+i,k_0;\underline{F})) \geq \Phi_{k_0+i}(\underline{V}(k_0+i),\underline{P}^o(k_0+i,k_0;\underline{F})) \\ &\geq \Phi_{k_0+i}(\underline{V}^o(k_0+i),\underline{P}^o(k_0+i,k_0;\underline{F})) = \underline{P}^o(k_0+i,k_0;\underline{F}) \quad . \end{aligned}\tag{2.5.13}$$

The theorem follows from induction.

<u>Definition 2.5.4</u>: The set of equations

$$\begin{aligned}\underline{P}(k+1,k_0;\underline{F}) &= (\underline{I}_n-\underline{V}(k)\underline{D}(k))(\underline{A}(k)\underline{P}(k,k_0;\underline{F})\underline{A}'(k)+\underline{Q}_1(k))(\underline{I}_n-\underline{V}(k)\underline{D}(k))' \\ &\quad +\underline{Q}_2(k)-\underline{Q}_1(k)+\underline{V}(k)\underline{R}(k)\underline{V}'(k) \quad ; \quad \underline{P}(k_0,k_0;\underline{F}) = \underline{F} \\ &\underline{V}(k)(\underline{R}(k)+\underline{D}(k)(\underline{A}(k)\underline{P}(k,k_0;\underline{F})\underline{A}'(k)+\underline{Q}_1(k))\underline{D}'(k)) \\ &\quad = (\underline{A}(k)\underline{P}(k,k_0;\underline{F})\underline{A}'(k) + \underline{Q}_1(k))\underline{D}'(k)\end{aligned}\tag{2.5.14}$$

is called the generalized Matrix Riccati Difference equation, and the unique solution is called the Riccati sequence, which is also the minimal sequence with respect to $\mathcal{B}_{k_0}$.

The above definition is meaningful because of theorem 2.5.3. In the special case when $\underline{R}(k)$ or $\underline{D}(k)\underline{Q}_1(k)\underline{D}'(k)$ (or both) is positive definite, then (2.5.14) can be written as a single nonlinear difference equation:

$$\underline{P}(k+1,k_0;\underline{F}) = \underline{A}(k)\underline{P}(k,k_0;\underline{F})\underline{A}'(k)+\underline{Q}_2(k)-[\underline{A}(k)\underline{P}(k,k_0;\underline{F})\underline{A}'(k)+\underline{Q}_1(k)]\underline{D}'(k)\cdot$$

$$\{\underline{R}(k)+\underline{D}(k)(\underline{A}(k)\underline{P}(k,k_0;\underline{F})\underline{A}'(k)+\underline{Q}_1(k))\underline{D}'(k)\}^{-1}\underline{D}(k)[\underline{A}(k)\underline{P}(k,k_0;\underline{F})\underline{A}'(k)+\underline{Q}_1(k)]$$

$$\underline{P}(k_0,k_0;\underline{F}) = \underline{F} \quad . \tag{2.5.15}$$

Equation (2.5.15) is the Matrix Riccati Difference equation.[28],[29]

2.6 Generalized Matrix Riccati Differential Equations

Let $\underline{V}(t)$ be arbitrary bounded measurable nxm matrix defined on $[t_0,T]$. Let $\underline{P}_v(t,t_0;\underline{F})$ be nxn matrix defined on $[t_0,T]$ which satisfies

$$\dot{\underline{P}}_v(t,t_0;\underline{F}) = (\underline{A}(t)-\underline{V}(t)\underline{D}_1(t))\underline{P}_v(t,t_0;\underline{F})+\underline{P}_v(t,t_0;\underline{F})(\underline{A}(t)-\underline{V}(t)\underline{D}_1(t))'$$

$$+\underline{V}(t)\underline{Q}(t)\underline{V}'(t)+(\underline{I}_n-\underline{V}(t)\underline{D}_2(t))\underline{R}(t)(\underline{I}_n-\underline{V}(t)\underline{D}_2(t))' \quad ;$$

$$\underline{P}_v(t_0,t_0;\underline{F}) = \underline{F} \geq \underline{0} \tag{2.6.1}$$

where $\underline{A}(t)$ is nxn, $\underline{D}_1(t)$, $\underline{D}_2(t)$ are mxn; $\underline{R}(t)$ is nonnegative definite nxn matrix and $\underline{Q}(t)$ is mxm nonnegative definite matrix (all matrices are assumed bounded measurable). Since (2.6.1) is linear, the solution $\underline{P}_v(t,t_0;\underline{F})$, $t \in [t_0,T]$, exists and is unique for a fixed bounded measurable $\underline{V}(t)$ (nxm) defined on $[t_0,T]$.

When $\underline{V}(t)$ ranges over the set of all bounded measurable mxm matrices defined on $[t_0,T]$, it generates the solution set $\mathcal{B}_{t_0}^T = \{\underline{P}_v(t,t_0;\underline{F}),$ $t \in [t_0,T] | \underline{V}(t)$ is bounded measurable nxm matrix defined on $[t_0,T]\}$.

Definition 2.6.1: (Minimal function) Let $\mathcal{S}_{t_0}^T = \{\underline{M}(t),\ t \in [t_0,T] | \underline{M}(t) \geq \underline{0},$ $t \in [t_0,T]\}$. An element $\underline{M}^o(\cdot) \in \mathcal{S}_{t_o}^T$ is called a minimal function with respect to $\mathcal{S}_{t_o}^T$ if for all $\underline{M}(\cdot) \in \mathcal{S}_{t_o}^T$, $\underline{M}^o(t) \leq \underline{M}(t)$, $t \in [t_0,T]$.

Let us define

$$\hat{\underline{A}}(t,V(t)) = \underline{A}(t) - \underline{V}(t)\underline{D}_1(t) \quad . \tag{2.6.2}$$

The solution of (2.6.1) is given by[34]

$$\underline{P}_v(t,t_0;\underline{F}) = \underline{\Phi}_{\hat{A}}(t,t_0)\underline{F}\ \underline{\Phi}'_{\hat{A}}(t,t_0) + \int_{t_0}^{t} \underline{\Phi}_{\hat{A}}(t,\tau)\{\underline{V}(\tau)\underline{Q}(\tau)\underline{V}'(\tau) +$$

$$[\underline{I}_n - \underline{V}(\tau)\underline{D}_2(\tau)]\cdot\underline{R}(\tau)[\underline{I}_n - \underline{V}(\tau)\underline{D}_2(\tau)]'\}\underline{\Phi}'_{\hat{A}}(t,\tau)d\tau \quad . \quad (2.6.3)$$

Since $\underline{F}$, $\underline{R}(\tau)$, $\underline{Q}(\tau)$ are all nonnegative definite matrices, we have

$$\underline{P}_v(t,t_0;\underline{F}) \geq \underline{0} \quad ; \quad t \in [t_0,T] \quad . \quad (2.6.4)$$

Define the matrix

$$\underline{\Psi}(t,\underline{V},\underline{P}) \triangleq \underline{\hat{A}}(t,\underline{V})\underline{P} + \underline{P}\ \underline{\hat{A}}'(t,\underline{V}) + \underline{V}\ \underline{Q}(t)\underline{V}' +$$

$$(\underline{I}_n - \underline{V}\ \underline{D}_2(t))\underline{R}(t)(\underline{I}_n - \underline{V}\ \underline{D}_2(t))' \quad (2.6.5)$$

where $\underline{V}$ is bounded nxm matrix, and $\underline{P}$ is bounded nxn matrix. For a fixed $\underline{P} \in M_{nn}$, define the set $\mathcal{V}_t(\underline{P}) = \{\underline{\tilde{V}} \in M_{nm} | (*)_t$ is satisfied where $(*)_t$ is the condition

$$(*)_t \qquad \underline{\tilde{V}}(\underline{Q}(t) + \underline{D}_2(t)\underline{R}(t)\underline{D}'_2(t)) = \underline{P}\ \underline{D}'_1(t) + \underline{R}(t)\underline{D}'_2(t) \quad .$$

<u>Lemma 2.6.2</u>: (Minimum Property) Let $\underline{P} \in M_{nn}$, and $\underline{P} \geq \underline{0}$; if $\underline{\tilde{V}} \in \mathcal{V}_t(\underline{P})$, then for all $\underline{V} \in M_{nm}$, we have

$$\underline{\Psi}(t,\underline{V},\underline{P}) \geq \underline{\Psi}(t,\underline{\tilde{V}},\underline{P}) \quad , \quad t \in [t_0,T] \quad . \quad (2.6.6)$$

<u>Proof</u>: Let $\underline{\tilde{V}} \in \mathcal{V}_t(\underline{P})$, by using $(*)_t$ we have

$$\underline{\Psi}(t,\underline{\tilde{V}},\underline{P})+(\underline{V}-\underline{\tilde{V}})(\underline{D}_2(t)\underline{R}(t)\underline{D}_2'(t)+\underline{Q}(t))(\underline{V}-\underline{\tilde{V}})'$$

$$= \underline{A}(t)\underline{P}-\underline{\tilde{V}}\ \underline{D}_1(t)\underline{P}+\underline{P}\ \underline{A}'(t)-\underline{P}\ \underline{D}_1'(t)\underline{\tilde{V}}'+\underline{R}(t)-\underline{\tilde{V}}\ \underline{D}_2(t)\underline{R}(t)-\underline{R}\ \underline{D}_2'(t)\underline{\tilde{V}}'$$

$$+2\underline{\tilde{V}}(\underline{D}_2(t)\underline{R}(t)\underline{D}_2'(t)+\underline{Q}(t))\underline{\tilde{V}}'+\underline{V}(\underline{D}_2(t)\underline{R}(t)\underline{D}_2'(t)+\underline{Q}(t))\underline{V}'-\underline{V}(\underline{D}_2(t)\underline{R}(t)\underline{D}_2'(t)$$

$$+\underline{Q}(t))\underline{\tilde{V}}'-\underline{\tilde{V}}(\underline{D}_2(t)\underline{R}(t)\underline{D}_2'(t)+\underline{Q}(t))\underline{V}'$$

$$= \underline{A}(t)\underline{P}-\underline{V}\ \underline{D}_1(t)\underline{P}+\underline{P}\ \underline{A}'(t)-\underline{P}\ \underline{D}_1'(t)\underline{V}'+\underline{R}(t)-\underline{V}\ \underline{D}_2(t)\underline{R}(t)-\underline{R}\ \underline{D}_2'(t)\underline{V}'$$

$$+\underline{V}(\underline{D}_2(t)\underline{R}(t)\underline{D}_2'(t)+\underline{Q}(t))\underline{V}' = \underline{\Psi}(t,\underline{V},\underline{P}) \qquad t \in [t_0,\tau] \quad . \qquad (2.6.7)$$

Since $\underline{R}(t) \geq \underline{0}$, $\underline{Q}(t) \geq \underline{0}$, (2.6.7) implies (2.6.6) immediately. From the lemma, we have

$$\underline{\Psi}(t,\underline{V}_1,\underline{P}) = \underline{\Psi}(t,\underline{V}_2,\underline{P}) \qquad \text{if} \qquad \underline{V}_1\underline{V}_2 \in \mathcal{V}_t(\underline{P}) \quad . \qquad (2.6.8)$$

Theorem 2.6.3: There exists a unique minimal function $\underline{P}^o(t,t_0;\underline{F})$, $t \in [t_0,T]$, with respect to the solution set $\mathcal{R}_{t_0}^T$.

Proof: Let us construct a sequence $\{\underline{P}_k(t,t_0;\underline{F})\}_{k=1}^{\infty}$ as follows: Set $\underline{P}_1(t,t_0;\underline{F}) = \underline{0}$, choose bounded measurable $\underline{V}_1(t) \in \mathcal{V}_t(\underline{P}_1(t,t_0;\underline{F})$ $t \in [t_0,T]$. Denote $\underline{P}_2(t,t_0;\underline{F}) = \underline{P}_{v_1}(t,t_0;\underline{F})$. Having chosen bounded measurable $\underline{V}_i(t) \in \mathcal{V}_t(\underline{P}_i(t,t_0;\underline{F}))$, $t \in [t_0,T]$, for $i = 1, \ldots, k$, let $\underline{P}_{k+1}(t,t_0;\underline{F}) = \underline{P}_{v_k}(t,t_0;\underline{F})$, $t \in [t_0,T]$. Using lemma 2.6.3, for $k > 1$:

$$\frac{d(\underline{P}_k(t,t_0;\underline{F})-\underline{P}_{k+1}(t,t_0;\underline{F}))}{dt} = \underline{\Psi}(t,\underline{V}_{k-1}(t),\underline{P}_k(t,t_0;\underline{F}))-\underline{\Psi}(t,\underline{V}_k(t),\underline{P}_{k+1}(t,t_0;\underline{F}))$$

$$\geq \underline{\Psi}(t,\underline{V}_k(t),\underline{P}_k(t,t_0;\underline{F}))-\underline{\Psi}(t,\underline{V}_k(t),\underline{P}_{k+1}(t,t_0;\underline{F}))$$

$$= \underline{\hat{A}}(t,\underline{V}_k(t))(\underline{P}_k(t,t_0;\underline{F})-\underline{P}_{k+1}(t,t_0;\underline{F}))+$$

$$(\underline{P}_k(t,t_0;\underline{F})-\underline{P}_{k+1}(t,t_0;\underline{F}))\underline{\hat{A}}(t,\underline{V}_k(t))' \quad . \qquad (2.6.9)$$

Since $\underline{P}_k(t_0,t_0;\underline{F}) = \underline{P}_{k+1}(t_0,t$ ... $= \ldots$ (2.6.9) implies that for $k > 1$:

$$\underline{P}_k(t,t_0;\underline{F}) \geq \underline{P}_{k+1} \ldots \geq \underline{0} \qquad t \in [t_0,T] \quad . \tag{2.6.10}$$

Therefore, there exists $\underline{P}^o(t,t_0;\underline{F})$ such that

$$\lim_{k\to\infty} \underline{P}_k(t,t_0;\underline{F}) = \underline{P}^o(t,t_0;\underline{F}) \quad . \tag{2.6.11}$$

Let us define, for $k > 1$, the matrix $\underline{P}^k(t,t_0;\underline{F})$ which satisfies

$$\dot{\underline{P}}^k(t,t_0;\underline{F}) = \underline{r}(t,\underline{V}_{k-1}(t),\underline{P}_{k-1}(t,t_0;\underline{F}));\underline{P}^k(t_0,t_0;\underline{F}) = \underline{F} \quad . \tag{2.6.12}$$

Clearly, $\underline{P}^k(t,t_0;\underline{F}) \geq \underline{0}$, $t \in [t_0,T]$, and

$$\frac{d(\underline{P}^k(t,t_0;\underline{F}) - \underline{P}_k(t,t_0;\underline{F}))}{dt} = \hat{\underline{A}}(t,\underline{V}_{k-1}(t)(\underline{P}_{k-1}(t,t_0;\underline{F}) - \underline{P}_k(t,t_0;F))$$
$$+ (\underline{P}_{k-1}(t,t_0;\underline{F}) - \underline{P}_k(t,t_0;\underline{F}))\hat{\underline{A}}'(t,\underline{V}_{k-1}(t)) \quad . \tag{2.6.13}$$

Since $\hat{\underline{A}}(t,\underline{V}_{k-1}(t))$ is bounded measurable in $[t_0,T]$, taking limits on both sides of (2.6.13) and using (2.6.11), (2.6.12) we have

$$\underline{r}(t,\underline{V}^o(t),\underline{P}^o(t,t_0;\underline{F})) = \lim_{k\to\infty} \dot{\underline{P}}^k(t,t_0;\underline{F}) = \lim_{k\to\infty} \dot{\underline{P}}_k(t,t_0;\underline{F}) = \dot{\underline{P}}^o(t,t_0;\underline{F}) \tag{2.6.14}$$

where $\underline{V}^o(t) \in \mathcal{V}_t(\underline{P}^o(t,t_0;\underline{F}))$, $t \in [t_0,T]$.

Note that the choice of the sequence $\{\underline{V}_i(t)\}_{i=1}^{\infty}$ is nonunique and so the sequence $\{\underline{P}_i(t,t_0;\underline{F})\}_{i=1}^{\infty}$ thus constructed is nonunique. Let $\{\tilde{\underline{V}}_i(t)\}_{i=1}^{\infty}$ be another chosen sequence where for $i \geq 1$, $\tilde{\underline{V}}_i(t) \in \mathcal{V}_t(\tilde{\underline{P}}_i(t,t_0;\underline{F}))$ and $\tilde{\underline{P}}_1(t,t_0;\underline{F}) = \underline{0}$, $\tilde{\underline{P}}_{i+1}(t,t_0;\underline{F}) = \underline{P}_{\tilde{V}_i}(t,t_0;\underline{F})$. Let

$$\lim_{k\to\infty} \tilde{\underline{P}}_k(t,t_0;\underline{F}) = \tilde{\underline{P}}^o(t,t_0;\underline{F}) \quad . \tag{2.6.15}$$

Then $\underline{\tilde{P}}^o(t,t_0;\underline{F})$ also satisfies

$$\dot{\underline{\tilde{P}}}^o(t,t_0;\underline{F}) = \underline{\Psi}(t,\underline{\tilde{V}}^o(t),\underline{\tilde{P}}^o(t,t_0;\underline{F})) \quad ; \quad \underline{\tilde{P}}^o(t_0,t_0;\underline{F}) = \underline{F} \qquad (2.6.16)$$

where $\underline{\tilde{V}}^o(t) \in \mathcal{U}_t(\underline{\tilde{P}}^o(t,t_0;\underline{F}))$, $t \in [t_0,T]$. Using lemma 2.6.2, we have

$$\begin{aligned}\dot{\underline{\tilde{P}}}^o(t,t_0;\underline{F}) - \dot{\underline{P}}^o(t,t_0;\underline{F}) &= \underline{\Psi}(t,\underline{\tilde{V}}^o(t),\underline{\tilde{P}}^o(t,t_0;\underline{F})) - \underline{\Psi}(t,\underline{V}^o(t),\underline{P}^o(t,t_0;\underline{F}))\\ &\geq \underline{\Psi}(t,\underline{V}^o(t),\underline{\tilde{P}}^o(t,t_0;\underline{F})) - \underline{\Psi}(t,\underline{V}^o(t),\underline{P}^o(t,t_0;\underline{F}))\\ &= \underline{\hat{A}}(t,\underline{V}^o(t))(\underline{\tilde{P}}^o(t,t_0;\underline{F}) - \underline{P}^o(t,t_0;\underline{F})) +\\ &\quad (\underline{\tilde{P}}^o(t,t_0;\underline{F}) - \underline{P}^o(t,t_0;\underline{F}))\underline{\hat{A}}'(t,\underline{V}^o(t)) \quad . \qquad (2.6.17)\end{aligned}$$

We conclude that $\underline{\tilde{P}}^o(t,t_0;\underline{F}) \geq \underline{P}^o(t,t_0;\underline{F})$, $t \in [t_0,T]$. We can interchange between $\underline{P}^o(t,t_0;\underline{F})$ and $\underline{\tilde{P}}^o(t,t_0;\underline{F})$ in (2.6.17) to obtain $\underline{\tilde{P}}^o(t,t_0;\underline{F}) \leq \underline{P}^o(t,t_0;\underline{F})$.

Therefore we have the uniqueness of the function $\underline{P}^o(t,t_0;\underline{F})$.

Let $\underline{V}(t)$ be an arbitrary bounded measurable nxm matrix. We have as before:

$$\begin{aligned}\dot{\underline{P}}_v(t,t_0;\underline{F}) - \dot{\underline{P}}^o(t,t_0;\underline{F}) &= \underline{\Psi}(t,\underline{V}(t),\underline{P}_v(t,t_0;\underline{F})) - \underline{\Psi}(t,\underline{V}^o(t),\underline{P}^o(t,t_0;\underline{F}))\\ &\geq \underline{\Psi}(t,\underline{V}(t),\underline{P}_v(t,t_0;\underline{F})) - \underline{\Psi}(t,\underline{V}(t),\underline{P}^o(t,t_0;\underline{F}))\\ &= \underline{\hat{A}}(t,\underline{V}(t))(\underline{P}_v(t,t_0;\underline{F}) - \underline{P}^o(t,t_0;\underline{F}))\\ &\quad + (\underline{P}_v(t,t_0;\underline{F}) - \underline{P}^o(t,t_0;\underline{F}))\underline{\hat{A}}'(t,\underline{V}(t)) \qquad (2.6.17)\end{aligned}$$

and so $\underline{P}_v(t,t_0;\underline{F}) \geq \underline{P}^o(t,t_0;\underline{F})$, $t \in [t_0,T]$. This completes the proof of the theorem. Note that the proof also gives an explicit algorithm to find $\underline{P}^o(t,t_0;\underline{F})$.

Definition 2.6.4: The set . . . .

$$\dot{\underline{P}}(t,t_0;\underline{F}) = (\underline{A}(t)-\underline{V}(t)\underline{D}_1(t)) \ldots ;\ldots)+\underline{P}(t,t_0;\underline{F})(\underline{A}(t)-\underline{V}(t)\underline{D}_1(t))'$$

$$+\underline{V}(t)\underline{Q}(t)\underline{V}'(t)+(\underline{I}_n-\underline{V}(t)\underline{D}_2(t))\underline{R}(t)(\underline{I}_n-\underline{V}(t)\underline{D}_2(t))' \quad ; \quad \underline{P}(t_0,t_0;\underline{F}) = \underline{F}$$

$$\underline{V}(t)(\underline{Q}(t)+\underline{D}_2(t)\underline{R}(t)\underline{D}_2'(t)) = \underline{P}(t,t_0;\underline{F})\underline{D}_1'(t)+\underline{R}(t)\underline{D}_2'(t) \qquad (2.6.18)$$

is called the generalized Matrix Riccati Differential Equation. The unique solution $\underline{P}(t,t_0;\underline{F})$, $t \in [t_0,T]$, is called the Riccati function, which is also the minimal function with respect to the solution set $\mathcal{B}^T_{t_0}$.

If $\underline{\Delta}(t) \triangleq \underline{Q}(t) + \underline{D}_2(t)\underline{R}(t)\underline{D}_2'(t) > \underline{0}$, then (2.6.18) reduces to a single nonlinear matrix differential equation:

$$\dot{\underline{P}}(t,t_0;\underline{F}) = (\underline{A}(t)-\underline{R}(t)\underline{D}_2'(t)\underline{\Delta}^{-1}(t)\underline{D}_1(t))\underline{P}(t,t_0;\underline{F})+\underline{P}(t,t_0;\underline{F})(\underline{A}(t)$$

$$-\underline{R}(t)\underline{D}_2'(t)\underline{\Delta}^{-1}(t)\underline{D}_1(t))'$$

$$-\underline{P}(t,t_0;\underline{F})\underline{D}_1(t)\underline{\Delta}^{-1}(t)\underline{D}_1'(t)\underline{P}(t,t_0;\underline{F})+\underline{R}(t)-\underline{R}(t)\underline{D}_2(t)\underline{\Delta}^{-1}(t)\underline{D}_2'(t)\underline{R}(t)$$

$$\underline{P}(t_0,t_0;\underline{F}) = \underline{F} \quad . \qquad (2.6.19)$$

Equation (2.6.19) is the Riccati Differential Equation.[31],[32]

In the general case, for a fixed bounded measurable $\underline{V}(t)$, $t \in [t_0,T]$:

$$\underline{P}_v(t,t_0;\underline{F}_1) \geq \underline{P}_v(t,t_0;\underline{F}_2) \quad \text{if} \quad \underline{F}_1 \geq \underline{F}_2 \quad . \qquad (2.6.20)$$

Let $\underline{V}_1(t) \in \mathcal{U}_t(\underline{P}(t,t_0;\underline{F}_1))$, $t \in [t_0,T]$, where $\underline{P}(t,t_0;\underline{F}_1)$ is the Riccati function satisfying (2.6.18). By theorem 2.6.3, we have for $\underline{F}_1 \geq \underline{F}_2$:

$$\underline{P}(t,t_0;\underline{F}_2) \leq \underline{P}_{v_1}(t,t_0;\underline{F}_2) \leq \underline{P}_{v_1}(t,t_0;\underline{F}_1) = \underline{P}(t,t_0;\underline{F}_1) \quad ; \quad t \in [t_0,T] \quad . \qquad (2.6.21)$$

2.7 Perspective

Measure theoretic approaches to probability theory can be found in Loeve [2], Doob [1]. The notion of statistics as used here was introduced by Halmas and Savage [13]. The term observation statistic is used so as to conform with physical interpretation. Conditional expectation and conditional distribution of a random variable (or vector) are treated in detail by Doob [1]. Conditional independence of sub-σ-algebra was treated by Meyer [14]. This is a more general and more intuitive definition of independence. Gaussian random vectors and Gaussian random processes are treated by Doob [1], Loeve [2], Cramer [15], Davenport and Root [3]. Gaussian white noise process viewed as the formal derivative of a Weiner process is treated by Wonham [4], McKean [16], Itô [17]; Gaussian white noise process viewed as a generalized process can be found in Tse [5], Gel'fand and Vilenkin [18].

Linear transformation of a Gaussian Vector is treated by Davenport and Root [3], Cramer [15], Doob [1]. Stochastic differential equations are studied by Itô [17], Stratonovich [19], Wong and Zakai [20], Tse [5], Clark [21]. Different interpretations to the stochastic differential equations are possible, some are in accordance with physical interpretation [19], [21] while some in terms of mathematical rigor.[17] In the linear case, all different interpretations are equivalent. The treatment used in Section 3 is consistent with all interpretations. The diffusion process is treated following Wonham [22] using Itô's interpretation, for a detailed discussion on the differential operator of a diffusion process, see Dynkin [7]. The proof of Itó's integration formula is given by Itô [17], Skorokhod [8], Wonham [4].

Matrix Riccati Difference equations are not treated in detail in the existing literature. Deyst and Price [28], Sorenson [29] and Aoki [30] considered the matrix Riccati difference equation which appears in filtering problems. Their considerations are restricted to a special, yet a large class of problems. The treatment given here is n , and the intrinsic properties of the matrix Riccati difference equation are revealed. The definition of generalized matrix Riccati difference equations and Riccati sequences are due to the author.In the continuous case, Kleinman [31], Wonham [32] had made detailed instigations. The approach used here is due to Wonham [32]. The generalization given in Section 6 is new, and the definition of generalized matrix Riccati differential equations and Riccati functions are due to the author. The motivation for this generalization is to bring out the most intrinsic properties of the equation and its solution. As we shall see in later chapters, this generalization allows us to understand the structural behavior of estimators and closed loop control systems.

CHAPTER III

OBSERVER THEORY FOR DISCRETE-TIME LINEAR SYSTEMS

3.1 Introduction

The problem of estimating the state variables of a dynamical system given observations of the output variables is of fundamental importance in the design of an optimal control system. If one considers the class of linear systems, then there are two approaches available in the literature. If the output variables can be measured exactly, and if there are no other stochastic disturbances acting on the system, then one can use a deterministic observer (see references [35], [36]). On the other hand, if all the output variables are corrupted by additive white noise, then one can use a Kalman filter (see references [39], [40], [37], [10]) for state estimation.

There are many cases in which some of the output variables are noise-free while others are noisy. One can argue that no measurement is exactly noise-free. On the other hand, there are many engineering systems in which the accuracy of measuring one variable is much greater than the accuracy of measuring some others. In such problems the measurement covariance matrix is almost singular and it can lead to ill-conditioned matrices and numerical problems. Thus, one can attempt to model the very accurate measurements as being deterministic.

The main purpose of this chapter is to examine this class of problems. In this contribution we examine the state estimation problem for linear discrete-time time-varying dynamical systems. The continuous time case will be considered in Chapter IV.

The structure of this chapter is as follows. In Section 3.2, we focus our attention to time-varying deterministic systems, we define the notion of a deterministic observer and estimator, the class of equivalent observers, and the class of minimal order observers. In Section 3.3, we extend the deterministic notion to the class of stochastic systems where we show that the class of equivalent observers yield unbiased estimates. Then we determine the class of observers that yield minimum variance estimates by formulating the problem as finding the minimal sequence of a certain solution set and then make use of theorem 2.5.3; we then prove that these observers yield indeed the conditional mean estimates of the state. Naturally, if the observation covariance matrix is positive definite one obtains the well-known Kalman filter. In Section 3.4, we examine in detail the case that some measurements are noisy while others are noise free. Under these conditions we show that the order of the minimal order observer is less than that of the state to be estimated. In Section 3.5, the notion of <u>detectability</u> is defined and the relation between detectability and observability of discrete linear system is considered; also in this section, we generalize the results of Kalman [41] on deadbeat deterministic observers to the time varying case. Using the concept of detectability, we derive necessary and sufficient condition for the minimum error covariance to be uniformly bounded and to have a steady state behavior. This is carried out in Section 3.6. In Section 3.7, we have general discussions on the approaches and results. In Section 3.8, detailed literature connected with the development in this chapter is listed.

3.2 Classes of Observers for Deterministic Systems

In this section we shall consider a linear time-varying discrete system S_1 described by

$$\left.\begin{array}{r} \text{(state eq.) } \underline{x}(k+1) = \underline{A}(k)\underline{x}(k) + \underline{B}(k)\underline{u}(k) \\ \text{(output eq.) } \underline{y}(k) = \underline{C}(k)\underline{x}(k) \end{array}\right. \quad S_1$$

where $k = 0, 1, 2, \ldots, \underline{x}(k) \in R^n$, $\underline{u}(k) \in R^r$, $\underline{y}(k) \in R^m$, and $\underline{C}(k)$ is of rank m.

Let M_{mn} be the set of all mxn matrices with real entries. If $m \leq n$, the null space of a matrix $\underline{M} \in M_{mn}$ will be denoted by $N(\underline{M}) = \{\underline{x} \in R^n; \underline{Mx} = \underline{0}_m \in R^m\}$.

Definition 3.2.1: Let $\underline{C}(k) \in M_{mn}$ be of rank m; the set

$$\Omega(\underline{C}(k);m,s,n) = \{\underline{T}(k) \in M_{sn}:N(\underline{T}(k)) \cap N(\underline{C}(k)) = \underline{0}_n \in R^n\}$$

is called the set of complimentary matrices of order s for $\underline{C}(k)$ if $s \geq n - m$.

We note that $\underline{T}(k) \in \Omega(\underline{C}(k);m,s,n)$ if and only if there exist $\underline{P}(k)$, $\underline{V}(k)$ of appropriate dimensions such that

$$\underline{P}(k)\underline{T}(k) + \underline{V}(k)\underline{C}(k) = \underline{I}_n \qquad (\underline{I}_n \in M_{nn}) \quad . \tag{3.2.1}$$

Definition 3.2.2: A discrete linear time varying system of dimension $s \geq n - m$ described by the relation

$$\Theta^1: \qquad \underline{z}(k+1) = \underline{F}(k)\underline{z}(k) + \underline{D}(k)\underline{y}(k) + \underline{G}(k)\underline{u}(k) \tag{3.2.2}$$

is called an s-order observer for the system S_1 if by some appropriate choice of $\underline{z}(0)$, we have

$$\underline{z}(k) = \underline{T}(k)\underline{x}(k) \quad ; \quad \text{for all } k = 0, 1, \ldots \tag{3.2.3}$$

for some $\underline{T}(k) \in \Omega(\underline{C}(k);m,s,n)$, $k = 0, 1, 2, \ldots$. We shall also say that

the observer is described by (3.2.3), and refer to such an observer by the symbol $\mathcal{O}_T^1$.

If $\mathcal{O}_T^1$ is an s-order observer for $\mathcal{S}_1$, then by an appropriate choice of $\underline{z}(0)$, we can reconstruct $\underline{x}(k)$ by

$$\underline{w}(k) = \underline{P}(k)\underline{z}(k) + \underline{V}(k)\underline{y}(k) = \underline{P}(k)\underline{T}(k)\underline{x}(k) + \underline{V}(k)\underline{C}(k)\underline{x}(k) = \underline{x}(k) \quad (3.2.4)$$

where $\underline{P}(k)$, $\underline{V}(k)$ are chosen to satisfy (3.2.1).

In the following theorem, we prove that a class of observers can be constructed for any linear discrete time varying system.

Theorem 3.2.3: Let $T = \{\underline{T}(k)\}_{k=0}^{\infty}$ be any sequence of matrices in M_{ns} such that $\underline{T}(k) \in \Omega(\underline{C}(k);m,s,n)$. Then, there exists an s-order observer, $\mathcal{O}_T^1$, for $\mathcal{S}_1$.

Proof: The proof is a constructive one in which an explicit form of $\mathcal{O}_T^1$ is obtained. Let $\underline{T}(k) \in \Omega(\underline{C}(k);m,s,n)$, k = 0, 1, ..., be given. Pick

$$\underline{F}(k) = \underline{T}(k+1)\underline{A}(k)\underline{P}(k) \quad (3.2.5)$$

$$\underline{D}(k) = \underline{T}(k+1)\underline{A}(k)\underline{V}(k) \quad (3.2.6)$$

$$\underline{G}(k) = \underline{T}(k+1)\underline{B}(k) \quad (3.2.7)$$

where $\underline{P}(k)$, $\underline{V}(k)$ satisfy (3.2.1), k = 0, 1, 2, ... Then

$$\underline{z}(k+1) - \underline{T}(k+1)\underline{x}(k+1) = \underline{T}(k+1)\underline{A}(k)\underline{P}(k)(\underline{z}(k) - \underline{T}(k)\underline{x}(k)) \quad . \quad (3.2.8)$$

Therefore, if we choose $\underline{z}(0) = \underline{T}(0)\underline{x}(0)$, we obtain

$$\underline{z}(k) = \underline{T}(k)\underline{x}(k) \quad ; \quad k = 0, 1, 2, \ldots \quad . \quad (3.2.9)$$

The observer described by the given sequence T has the explicit form:

$$\mathcal{O}_T^1 : \underline{z}(k+1) = \underline{T}(k+1)\underline{A}(k)\underline{P}(k)\underline{z}(k) + \underline{T}(k+1)\underline{A}(k)\underline{V}(k)\underline{y}(k) + \underline{T}(k+1)\underline{B}(k)\underline{u}(k) \quad . \tag{3.2.10}$$

To an observer $\mathcal{O}_T^1$, we associate an estimator $\mathcal{E}_T^1$ described by (see Figure 3.1)

$$\mathcal{E}_T^1 : \begin{cases} \underline{z}(k+1) = \underline{T}(k+1)\underline{A}(k)\underline{P}(k)\underline{z}(k) + \underline{T}(k+1)\underline{A}(k)\underline{V}(k)\underline{y}(k) + \underline{T}(k+1)\underline{B}(k)\underline{u}(k) \\ \underline{w}(k) = \underline{P}(k)\underline{z}(k) + \underline{V}(k)\underline{y}(k) \end{cases} \tag{3.2.11}$$

where $\underline{P}(k)$, $\underline{V}(k)$ satisfy (3.2.1) for the fixed $\underline{T}(k)$, k = 0, 1, ... By setting $\underline{z}(0) = \underline{T}(0)\underline{x}(0)$, $\underline{w}(k)$ will equal $\underline{x}(k)$ by (3.2.4). But in most cases, the initial state $\underline{x}(0)$ is unknown. We shall fix the initial condition for the observer $\mathcal{O}_T^1$ by the relation

$$\underline{z}(0) = \underline{T}(0)\underline{\alpha} \tag{3.2.12}$$

where the vector $\underline{\alpha}$ is a guess for $\underline{x}(0)$. Thus $\underline{\alpha}$ is any vector in R^n, and the possible values of $\underline{z}(0)$ will be in the range space of $\underline{T}(0)$.

Let $V \triangleq \{\underline{V}(k)\}_{k=0}^{\infty}$ be any sequence of matrices in M_{nm}. Let us associate with the given sequence a sequence of sets where

$$\mathcal{T}_{\underline{V}(k)} = \{\underline{T}(k)\varepsilon M_{sn} \mid \underline{P}(k)\underline{T}(k)+\underline{V}(k)\underline{C}(k) = \underline{I}_n \text{ for some } \underline{P}(k)\varepsilon M_{ns}; s\geq n-m\} \quad ,$$

$$k = 0, 1, \ldots \quad .$$

If $T = \{\underline{T}(k)\}_{k=0}^{\infty}$ is a sequence of matrices in M_{sn} such that $\underline{T}(k) \varepsilon \mathcal{T}_{\underline{V}(k)}$, k = 0, 1, 2, ..., then we shall in short write $T \varepsilon \mathcal{T}_V$. Now let $T \varepsilon \mathcal{T}_V$; by theorem 3.2.3 we can associate to every such T with an

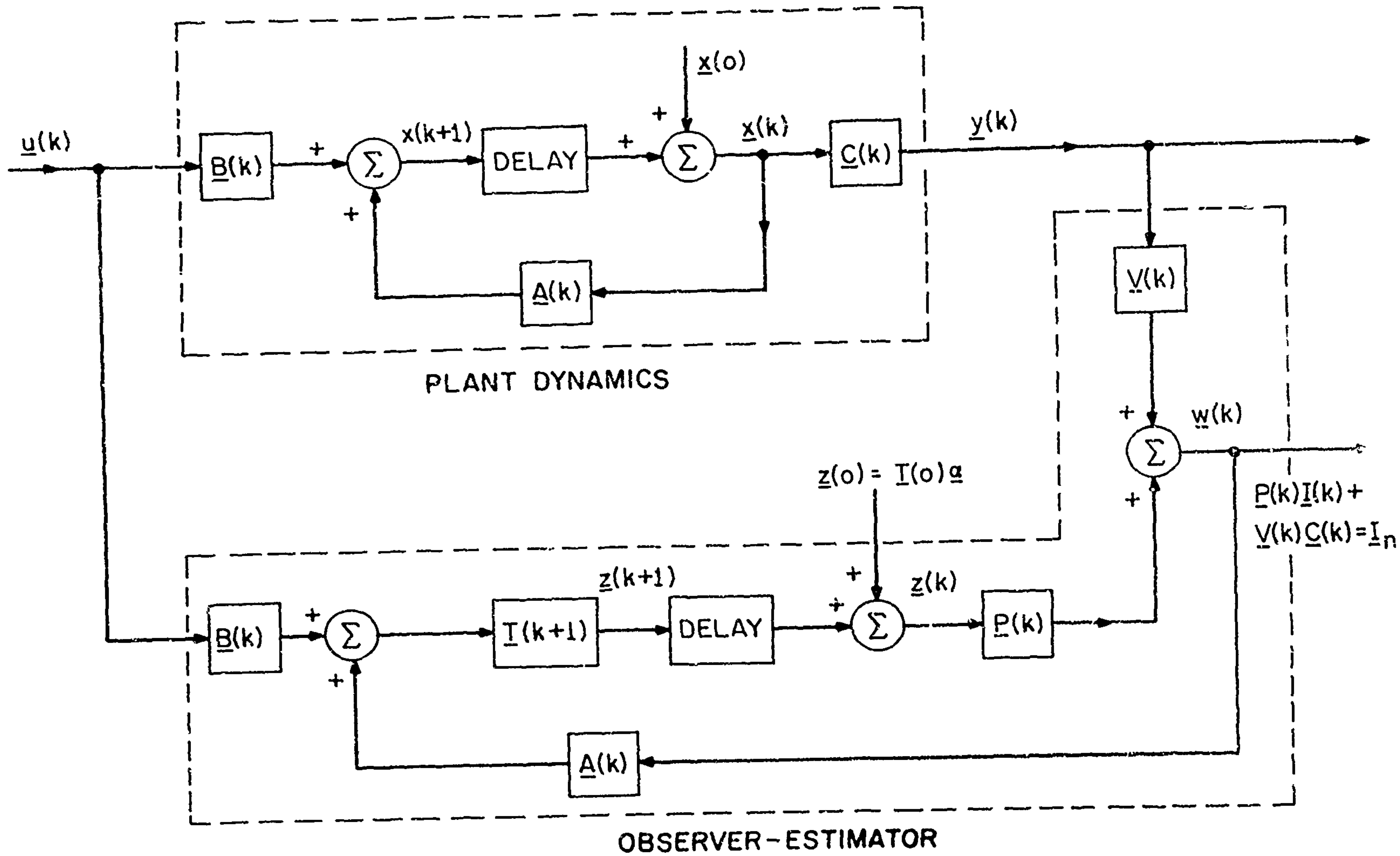


Fig. 3.1 THE STRUCTURE OF AN OBSERVER-ESTIMATOR FOR A DISCRETE TIME DETERMINISTIC LINEAR SYSTEM. THE VECTOR $\underline{w}(k)$ IS THE ESTIMATE OF THE STATE $\underline{x}(k)$.

observer Θ_T^1, and an estimator $\mathcal{E}_T^1$. Thus, we can associate with any fixed V a class of observers of different orders.

Suppose that for a fixed V, the sequence of matrices $\{(\underline{I}_n - \underline{V}(k)C(k))\}$, k = 0, 1, 2, ..., has rank n − p; then the class $\Pi_V^p = \{\Theta_T^1 | T \in \mathfrak{T}_V$, and $\underline{T}(k) \in M_{n(n-p)}$ has full rank, k = 0, 1, ...} is called the class of <u>minimal order observers associated</u> with V.

<u>Definition 3.2.4</u>: Let L_1 be a linear discrete system described by

$$L_1: \begin{cases} \underline{x}_1(k+1) = \underline{F}_1(k)\underline{x}_1(k) + \underline{G}_1(k)\underline{u}(k) \\ \underline{y}_1(k) = \underline{C}_1(k)\underline{x}_1(k) + \underline{D}_1(k)\underline{u}(k) \end{cases} \quad (3.2.13)$$

with $\underline{x}_1(0) \in X_1 \subset R^n$. We shall say that L_2, described by

$$L_2: \begin{cases} \underline{x}_2(k+1) = \underline{F}_2(k)\underline{x}_2(k) + \underline{G}_2(k)\underline{u}(k) \\ \underline{y}_2(k) = \underline{C}_2(k)\underline{x}_2(k) + \underline{D}_2(k)\underline{u}(k) \end{cases} \quad (3.2.14)$$

with $\underline{x}_2(0) \in X_2 \subset R^n$, is <u>an equivalent representation</u> of L_1 if for any $\underline{\beta}_1 \in X_1$, there exists a $\underline{\beta}_2 \in X_2$ such that

$$\underline{\phi}_1(k;\underline{\beta}_1,\underline{u}_k) = \underline{\phi}_2(k;\underline{\beta}_2,\underline{u}_k) \quad \forall\, \underline{u}_k = \{\underline{u}(i)\}_{i=0}^k \quad (3.2.15)$$

where $\underline{\phi}_i(k;\underline{\beta}_i,\underline{u}_k)$ is the output of the system L_i for $\underline{x}_i(0) = \underline{\beta}_i$ and applied control $\underline{u}_k$.

Equivalent representations may not be symmetric, i.e., if L_2 is an equivalent representation of L_1, this does not imply that L_1 is an equivalent representation of L_2. If L_1 and L_2 are both equivalent representations of each other, then we say that L_1 and L_2 are <u>equivalent</u>.

We shall say that two [illegible], $\mathcal{O}^1_{\tilde{T}}$, are equivalent if $\mathcal{E}^1_T$ and $\mathcal{E}^1_{\tilde{T}}$ are equivalent. Let $\mathcal{O}^1_T$ be [illegible]-order observer; by some linear transformation on the state, we can easily see that we can construct an s'-order observer $\mathcal{O}^1_{\tilde{T}}$, which is an equivalent representation of $\mathcal{O}^1_T$, where s' [illegible] s. Interestingly enough, if we restrict the possible initial conditions of the observer state, then for any s-order observer, $\mathcal{O}^1_T$, we can find an s'-order observer, $\mathcal{O}^1_{\tilde{T}}$, which is an equivalent representation of $\mathcal{O}_T$, with s' $\geq$ s.

<u>Theorem 3.2.5</u>: Let $V = \{\underline{V}(k)\}_{k=0}^{\infty}$ be a fixed sequence of matrices in M_{mm}, the class of observers $\mathcal{O}^1_T$, $T \in \mathcal{T}_V$, are equivalent.

<u>Proof</u>: Let $\mathcal{E}^1_T$, $\mathcal{E}^1_{\tilde{T}}$ be any two estimators, $T, \tilde{T} \in \mathcal{T}_V$, described by

$$\mathcal{E}^1_T:\quad \begin{aligned} &\underline{z}(k+1) = \underline{T}(k+1)\underline{A}(k)\underline{P}(k)\underline{z}(k) + \underline{T}(k+1)\underline{A}(k)\underline{V}(k)\underline{y}(k) + \underline{T}(k+1)\underline{B}(k)\underline{u}(k) \\ &\underline{w}(k) = \underline{P}(k)\underline{z}(k) + \underline{V}(k)\underline{y}(k) \quad ; \quad \underline{z}(0) \in S = \{\underline{T}(0)\underline{\alpha} \mid \underline{\alpha} \in R^n\} \end{aligned} \tag{3.2.16}$$

$$\mathcal{E}^1_{\tilde{T}}:\quad \begin{aligned} &\underline{\tilde{z}}(k+1) = \underline{\tilde{T}}(k+1)\underline{A}(k)\underline{\tilde{P}}(k)\underline{\tilde{z}}(k) + \underline{\tilde{T}}(k+1)\underline{A}(k)\underline{V}(k)\underline{y}(k) + \underline{\tilde{T}}(k+1)\underline{B}(k)\underline{u}(k) \\ &\underline{\tilde{w}}(k) = \underline{\tilde{P}}(k)\underline{\tilde{z}}(k) + \underline{V}(k)\underline{y}(k) \quad ; \quad \underline{\tilde{z}}(0) \in \tilde{S} = \{\underline{\tilde{T}}(0)\underline{\alpha} \mid \underline{\alpha} \in R^n\} \end{aligned} \tag{3.2.17}$$

Let $\underline{z}(0) = \underline{T}(0)\underline{\alpha}_1 \in S$ be the initial condition for $\mathcal{O}^1_T$. Choose $\underline{\tilde{z}}(0) = \underline{\tilde{T}}(0)\underline{\alpha}_1 \in \tilde{S}$ be the initial condition for $\mathcal{O}^1_{\tilde{T}}$, then for all $\underline{y}(0)$:

$$\underline{w}(0) = \underline{P}(0)\underline{T}(0)\underline{\alpha}_1 + \underline{V}(0)\underline{y}(0) = \underline{\tilde{P}}(0)\underline{\tilde{T}}(0)\underline{\alpha}_1 + \underline{V}(0)\underline{y}(0) = \underline{\tilde{w}}(0) \quad . \tag{3.2.18}$$

Assume that $\underline{w}(k-1) = \underline{\tilde{w}}(k-1)$, then

$$\begin{aligned} \underline{w}(k) &= \underline{P}(k)\underline{T}(k)\underline{A}(k)\underline{w}(k-1) + \underline{V}(k)\underline{y}(k) + \underline{P}(k)\underline{T}(k)\underline{B}(k)\underline{u}(k) \\ &= \underline{\tilde{P}}(k)\underline{\tilde{T}}(k)\underline{A}(k)\underline{w}(k-1) + \underline{V}(k)\underline{y}(k) + \underline{\tilde{P}}(k)\underline{\tilde{T}}(k)\underline{B}(k)\underline{u}(k) \\ &= \underline{\tilde{w}}(k) \quad . \end{aligned}$$

From (3.2.18) and (3.2.19) we conclude that

$$\underline{w}(k) = \underline{\tilde{w}}(k) \qquad \text{for all } k \quad . \tag{3.2.20}$$

Conversely if $\underline{\tilde{z}}(0) = \underline{T}(0)\underline{\tilde{s}}_2 \in \tilde{S}$, pick $\underline{z}(0) = \underline{T}(0)\underline{s}_2 \in S$, then we also have that the output of $\mathcal{S}^1_T$ and that of $\mathcal{S}^1_{\tilde{T}}$ are the same. Thus $\mathcal{S}^1_T$, $\mathcal{S}^1_{\tilde{T}}$ are equivalent, and $\mathcal{G}^1_T$, $\mathcal{G}^1_{\tilde{T}}$ are equivalent.

Thus, for a fixed sequence V, we can associate with a class of equivalent observers $\mathcal{G}^1_T$, $T \in \tilde{\mathcal{J}}_V$. When V ranges over all possible sequences, we obtain different classes of observers parameterized by the sequence V. In a vague sense, the class of observers $\mathcal{G}^1_T$, $T \in \tilde{\mathcal{J}}_V$, utilize the same amount of incoming information provided by the observations $\underline{y}(k)$, $k = 0, 1, \ldots$. The notion of efficiency of a system, as regard to the processing of incoming information, is (in a loose sense) a ratio between information utilized and the complexity of the system. Thus for a fixed V, the most efficient system associated with it is the class of minimal order observers, π^p_V. In view of the above discussion, the design of appropriate observer for estimation and control purposes can be split into two distinct steps: 1) to find the appropriate V^* which specifies the operating performance of the class of observers, 2) to find an observer in the class of minimal order observers, $\pi^p_{V^*}$.

3.3 Optimum Classes of Observers for Linear Stochastic Systems

Let us consider the stochastic system $\mathcal{S}_2$ described by

$$\mathcal{S}_2: \quad \begin{aligned} \underline{x}(k+1) &= \underline{A}(k)\underline{x}(k) + \underline{B}(k)\underline{u}(k) + \underline{\xi}(k) \\ \underline{y}(k) &= \underline{C}(k)\underline{x}(k) + \underline{\eta}(k) \end{aligned} \tag{3.3.1}$$

where $\underline{x}(0)$, $\underline{\xi}(0)$, $\underline{n}(0)$, $\underline{\xi}(1)$, ... are independent Gaussian random vectors with statistical law.

$$\underline{x}(0) \sim G(\underline{x}_0, \underline{\Sigma}_0) \quad ; \quad \underline{\Sigma}_0 > \underline{0} \tag{3.3.2}$$

$$\underline{\xi}(k) \sim G(\underline{0}, \underline{R}(k)) \quad ; \quad \underline{R}(k) \geq \underline{0} \tag{3.3.3}$$

$$\underline{n}(k) \sim G(\underline{0}, \underline{Q}(k)) \quad ; \quad \underline{Q}(k) \geq \underline{0} \quad . \tag{3.3.4}$$

The control $\underline{u}(k)$, $k = 0, 1, \ldots$, is an arbitrary but known sequence.

Let $V = \{\underline{V}(k)\}_{k=0}^{\infty}$ be an arbitrary sequence of matrices in M_{nm}. If we use an estimator $\mathcal{E}_T^1$, $T \in \mathcal{J}_V$, for $\mathcal{S}_2$ to generate an estimate of $\underline{x}(k)$, then the error $\underline{e}(k) \triangleq \underline{w}(k) - \underline{x}(k)$, can be computed from (3.2.11) and (3.3.1). By picking $\underline{z}(0) = \underline{T}(0)\underline{x}_o$, the error dynamics are given by

$$\begin{aligned}\underline{e}(k+1) &= [\underline{I}_n-\underline{V}(k+1)\underline{C}(k+1)]\underline{A}(k)\underline{e}(k)+\underline{V}(k+1)\underline{n}(k+1)+(\underline{V}(k+1)\underline{C}(k+1)-\underline{I}_n)\underline{\xi}(k)\\ \underline{e}(0) &= [\underline{I}_n-\underline{V}(0)\underline{C}(0)][\underline{x}_o-\underline{x}(0)]+\underline{V}(0)\underline{n}(0)\end{aligned} \tag{3.3.5}$$

So explicitly (3.3.5) reveals that all estimators $\mathcal{E}_T^1$, $T \in \mathcal{J}_V$, give the same error dynamics, which in some sense reflect the state of uncertainty of the system $\mathcal{S}_2$. From (3.3.5) we see that

$$E[\underline{e}(k)] = \underline{0} \quad ; \quad k = 0, 1, 2, \ldots \quad . \tag{3.3.6}$$

Therefore, associated with an arbitrary V, we have a class of equivalent observers whose associated estimators yield unbiased estimates. Our aim now is to find the optimum V^* which will result in minimum error covariance. From (3.3.2) to (3.3.6) we see that the error covariance will propagate according to the matrix difference equation:

$$\underline{\Sigma}(k+1) = [\underline{I}_n-\underline{V}(k+1)\underline{C}(k+1)][\underline{A}(k)\underline{\Sigma}(k)\underline{A}'(k)+\underline{R}(k)][\underline{I}_n-\underline{V}(k+1)\underline{C}(k+1)]'$$
$$+\underline{V}(k+1)\underline{Q}(k+1)\underline{V}'(k+1) \quad ; \quad k = 0, 1, \ldots$$
$$\underline{\Sigma}(0) = [\underline{I}_n-\underline{V}(0)\underline{C}(0)]\underline{\Sigma}_o[\underline{I}_n-\underline{V}(0)\underline{C}(0)]'+\underline{V}(0)\underline{Q}(0)\underline{V}'(0) \tag{3.3.7}$$

where

$$\underline{\Sigma}(k) = E\{\underline{e}(k)\underline{e}'(k)\} \quad ; \quad k = 0, 1, \ldots \quad . \tag{3.3.8}$$

Defining

$$\underline{A}(-1) = \underline{I}_n \quad ; \quad \underline{R}(-1) = \underline{0} \tag{3.3.9}$$

equation (3.3.7) can be written

$$\underline{\Sigma}(k+1) = [\underline{I}_n-\underline{V}(k+1)\underline{C}(k+1)][\underline{A}(k)\underline{\Sigma}(k)\underline{A}'(k)+\underline{R}(k)][\underline{I}_n-\underline{V}(k+1)\underline{C}(k+1)]'$$
$$+\underline{V}(k+1)\underline{Q}(k+1)\underline{V}'(k+1); \quad k = -1, 0, 1, \ldots$$
$$\underline{\Sigma}(-1) = \underline{\Sigma}_o \tag{3.3.10}$$

When V ranges over all possible nxm matrices, we generate a solution set $\mathcal{B}_{-1}$ of (3.3.10). For the optimum estimation, we would like to find a sequence V^* which will give rise to the minimal sequence with respect to the solution set $\mathcal{B}_{-1}$. By comparing (3.3.10) with (2.5.1), we have the following:

Theorem 3.3.1: A unique minimal sequence $\{\underline{\Sigma}^*(k)\}_{k=0}^{\infty}$ with respect to the solution set $\mathcal{B}_{-1}$ of (3.3.9) exists and is given by

$$\underline{\Sigma}^*(k + 1) = \underline{\Delta}^*(k) - \underline{\tilde{V}}^*(k + 1)\underline{C}(k + 1)\underline{\Delta}^*(k) \quad ; \quad k = 0, 1, \ldots$$
$$\underline{\Sigma}^*(0) = \underline{\Sigma}_o - \underline{\Sigma}_o\underline{C}'(0)[\underline{C}(0)\underline{\Sigma}_o\underline{C}'(0) + \underline{Q}(0)]^{-1}\underline{C}(0)\underline{\Sigma}_o \tag{3.3.11}$$

where

$$\underline{\Delta}^*(k) \triangleq \underline{A}(k)\underline{\Sigma}^*(k)\underline{A}'(k) - \underline{B}(k) \quad ; \quad k = 0, 1, \ldots \qquad (3.3.12)$$

and $\tilde{\underline{V}}^*(k) \in \mathcal{V}_{k-1}(\underline{\Sigma}^*(k-1)) = \{\underline{V} \in M_{nm} : \underline{V}[\underline{C}(k)\underline{\Delta}^*(k-1)\underline{C}'(k)+\underline{Q}(k)] = \underline{\Delta}^*(k-1)\underline{C}'(k)\}$,

$k = 1, 2, \ldots$.

If either $\underline{Q}(k) > \underline{0}$, $k = 0, 1, \ldots$, or $\underline{C}(k+1)\underline{R}(k)\underline{C}'(k+1) > \underline{0}$, $k = 0, 1, \ldots$ (or both), then the <u>unique</u> Riccati sequence is given by

$$\underline{\Sigma}^*(k+1) = \underline{\Delta}^*(k)-\underline{\Delta}^*(k)\underline{C}'(k+1)[\underline{C}(k+1)\underline{\Delta}^*(k)\underline{C}'(k+1)+\underline{Q}(k+1)]^{-1}\underline{C}(k+1)\underline{\Delta}^*(k)$$

$$\underline{\Sigma}^*(0) = \underline{\Sigma}_o-\underline{\Sigma}_o\underline{C}'(0)[\underline{C}(0)\underline{\Sigma}_o\underline{C}'(0)+\underline{Q}(0)]^{-1}\underline{C}(0)\underline{\Sigma}_o \quad ; \quad k = 0, 1, \ldots \qquad (3.3.13)$$

where $\underline{\Delta}^*(k)$ is given by (3.3.12), and the unique $\{\underline{V}^*(k)\}_{k=0}^{\infty}$ which gives rise to the Riccati minimal sequence is given by

$$\underline{V}^*(k) = \underline{\Delta}^*(k-1)\underline{C}'(k)[\underline{C}(k)\underline{\Delta}^*(k-1)\underline{C}'(k) + \underline{Q}(k)]^{-1} \quad ; \quad k = 0, 1, \ldots . \qquad (3.3.14)$$

The proof of this theorem follows from theorem 2.5.3 directly by identifying

$$\underline{C}(k+1) \rightarrow \underline{D}(k) \quad , \quad \underline{R}(k) \rightarrow \underline{Q}_1(k) = \underline{Q}_2(k) \qquad (3.3.15)$$

$$\underline{\Sigma}(k) \rightarrow \underline{P}_V(k,-1;\underline{\Sigma}_o) \quad , \quad \underline{Q}(k) \rightarrow \underline{R}(k) \quad . \qquad (3.3.16)$$

Theorem 3.3.1 implies that <u>an</u> optimum class of observers is specified by <u>any</u> sequence $\{\tilde{\underline{V}}^*(k)\}_{k=0}^{\infty}$ where $\tilde{\underline{V}}^*(k) \in \mathcal{V}_{k-1}(\underline{\Sigma}^*(k-1))$ inductively, $k = 0, 1, \ldots$ with $\underline{\Sigma}^*(k)$ given by (3.3.11), and (3.3.12), $k = -1, 0, 1, \ldots$. In the special case when $\underline{Q}(k) > \underline{0}$, or $\underline{C}(k+1)\underline{R}(k)\underline{C}'(k+1) > \underline{0}$, $k = 0, 1, \ldots$, then there is a unique class of optimum observers specified by $\{\underline{V}^*(k)\}_{k=0}^{\infty}$, given by (3.3.11). In fact, one can show that an observer with an initial

condition $\underline{z}(0) = \underline{T}(0)\underline{x}_o$ is in some sense equivalent to the concept of unbiased linear estimator (see Section 3.7), and thus an optimum class of observers is also an optimum class of linear unbiased estimators. In the rest of this section, we shall show that when $\underline{u}(k)$ is known, the estimator generated via an observer $\mathcal{G}_T^2$, $T \in \mathcal{J}_{\bar{V}*}$,† and its associated estimator $\mathcal{E}_T^2$ is the conditional mean of $\underline{x}(k)$. This reflects the truly optimum nature of the optimum classes of observers.

Since $\underline{u}(k)$, $k = 0, 1, \ldots$, are known a priori, we may assume them to be zero without loosing generality. Now consider $\mathcal{S}_2$ with control sequence equal to zero. By the Gaussian assumption, the conditional expectation of $\underline{x}(k)$, denoted by

$$\hat{\underline{x}}(k|k) = E\{\underline{x}(k)|F(k)\} \quad ; \quad F(k) \triangleq F(\underline{y}(i),\ i = 0, 1, \ldots, k) \qquad (3.3.17)$$

equals almost surely to some linear functional of $\{\underline{y}(0), \ldots, \underline{y}(k)\}$.[1]

Lemma 3.3.2: (Weiner-Hopf Equation) Let $\{\underline{w}(k)\}_{k=0}^{\infty}$ be a sequence of random vectors such that $\underline{w}(k)$ is a linear functional of $\underline{y}(0), \ldots, \underline{y}(k)$. If in addition, $\underline{w}(k)$ satisfies for $k = 0, 1, \ldots$

$$E[\underline{w}(k)\underline{y}'(i)] = E[\underline{x}(k)\underline{y}'(i)] \qquad i = 0, 1, \ldots, k \qquad (3.3.18)$$

then $\underline{w}(k) = \hat{\underline{x}}(k|k)$ a.s. for all k.

The proof is given in the Appendix B. An immediate consequence of this lemma is the Projection Theorem.

---

†The superscript is used to indicate that the stochastic system $\mathcal{S}_2$ is being considered.

Theorem 3.3.3: (Projection Theorem) Let $\{\underline{w}(k)\}_{k=0}^{\infty}$ be a sequence of random vectors such that $\underline{w}(k)$ is a linear functional of $\{\underline{y}(0), \ldots, \underline{y}(k)\}$. Let $\underline{e}(k) \triangleq \underline{w}(k) - \underline{x}(k)$, $k = 0, 1, \ldots$. If $\underline{e}(k)$ satisfies, for all k,

$$E[\underline{e}(k)\underline{y}'(i)] = \underline{0} \quad ; \quad i = 0, 1, \ldots, k \tag{3.3.19}$$

then $\underline{w}(k) = \hat{\underline{x}}(k|k)$ almost surely for all k. [We shall refer to Equation (3.3.19) as the Projection equation.]

For any fixed sequence $V = \{\underline{V}(k)\}_{k=0}^{\infty}$; the output of the estimator $\mathcal{E}_T^2$, $T \in \mathcal{T}_V$, at time k is clearly a linear functional of $\{\underline{y}(0), \ldots, \underline{y}(k)\}$. In the following, we shall prove that

$$E\{\tilde{\underline{e}}^*(k)\underline{y}'(i)\} = \underline{0} \quad ; \quad i = 0, 1, \ldots, k \tag{3.3.20}$$

where $\tilde{\underline{e}}^*(k)$ the error of estimates if we adopt $\mathcal{E}_T^2$, $T \in T_{\tilde{V}^*}$, as an estimating device, and $\tilde{\underline{e}}^*(k)$ is given by (see 3.3.5)

$$\begin{aligned} \tilde{\underline{e}}^*(k+1) &= [\underline{I}_n - \tilde{\underline{V}}^*(k+1)\underline{C}(k+1)][\underline{A}(k)\tilde{\underline{e}}^*(k) - \underline{\xi}(k)] + \tilde{\underline{V}}^*(k+1)\underline{n}(k+1) \\ \tilde{\underline{e}}^*(0) &= [\underline{I}_n - \tilde{\underline{V}}^*(0)\underline{C}(0)](\underline{x}_o - \underline{x}(0)) + \tilde{\underline{V}}^*(0)\underline{n}(0) \end{aligned} \tag{3.3.21}$$

and $\tilde{\underline{V}}^*(k) \in \mathcal{U}_{k-1}(\underline{\Sigma}^*(k-1))$ inductively, $k = 0, 1, \ldots$, with $\underline{\Sigma}^*(k)$ given by (3.3.11), (3.3.12). Let us first establish a lemma and a corollary which will be useful in later discussions.

Lemma 3.3.4: Let $\{\tilde{\underline{e}}(k)\}_{k=0}^{\infty}$ be a sequence of random vectors satisfying (3.3.20). Let $\{\underline{x}(k)\}_{k=0}^{\infty}$ be given by (3.3.1) with $\underline{u}(k) = \underline{0}$, $k = 0, 1, \ldots$, then for all $k = 0, 1, \ldots$

$$E\{\tilde{\underline{e}}^*(k)\underline{x}'(k)\} = -\underline{\Sigma}^*(k) \tag{3.3.22}$$

where $\underline{\Sigma}^*(k)$ is given by (3.3.11), (3.3.12).

Proof: We shall use induction on k. For k = 0, using (3.3.21), (3.3.1), (3.3.11) and the given statistical law:

$$\begin{aligned} E\{\underline{\tilde{e}}^*(0)\underline{x}'(0)\} &= (\underline{I}_n - \underline{\tilde{V}}^*(0)\underline{C}(0))\{\underline{x}_o\underline{x}_o' - E[\underline{x}(0)\underline{x}'(0)]\} \\ &= -\underline{\Sigma}_o + \underline{\tilde{V}}^*(0)\underline{C}(0)\underline{\Sigma}_o = -\underline{\Sigma}^*(0) \quad . \end{aligned} \tag{3.3.23}$$

Assume that (3.3.22) is true for k = 0, 1, ..., n. Using (3.3.21), (3.3.1), (3.3.11), (3.3.12) and the given statistical law, we have:

$$\begin{aligned} E[\underline{\tilde{e}}^*(n+1)\underline{x}'(n+1)] &= [\underline{I}_n-\underline{\tilde{V}}^*(n+1)\underline{C}(n+1)][\underline{A}(n)E\{\underline{\tilde{e}}^*(n)\underline{x}'(n)\}\underline{A}'(n)-\underline{R}(n)] \\ &= -[\underline{\tilde{\Delta}}^*(n)-\underline{\tilde{V}}^*(n+1)\underline{C}(n+1)\underline{\tilde{\Delta}}^*(n)] = -\underline{\tilde{\Sigma}}^*(n+1) \quad . \end{aligned} \tag{3.3.24}$$

The lemma is proved by induction.

Corollary 3.3.5: Let $\{\underline{\tilde{e}}(k)\}_{k=0}^{\infty}$ be a sequence of random vectors satisfying (3.3.21) where $\underline{\tilde{V}}^*(k) \in \mathcal{V}_{k-1}(\underline{\Sigma}^*(k-1))$ with $\underline{\Sigma}^*(k-1)$ given by (3.3.11) and (3.3.12). Let $\{\underline{y}(k)\}_{k=0}^{\infty}$ be given by (3.3.1) with $\underline{u}(k) = \underline{0}$, k = 0, 1, ... Then for all k = 0, 1, ...

$$E\{\underline{\tilde{e}}^*(k)\underline{y}'(k)\} = \underline{0} \quad . \tag{3.3.25}$$

Proof: We shall use induction on k. For k = 0, since using (3.3.21), (3.3.1), (3.3.22) and the given statistical law, we have

$$\begin{aligned} E\{\underline{\tilde{e}}^*(0)\underline{y}'(0)\} &= E\{\underline{\tilde{e}}^*(0)\underline{x}'(0)\}\underline{C}'(0)+E\{\underline{\tilde{e}}^*(0)\underline{n}'(0)\} \\ &= -\underline{\Sigma}_o\underline{C}'(0)+\underline{\tilde{V}}^*(0)\underline{C}(0)\underline{\Sigma}_o\underline{C}'(0)+\underline{\tilde{V}}^*(0)\underline{Q}(0) = \underline{0} \quad . \end{aligned} \tag{3.3.26}$$

Assume that (3.3.25) is true for k = 0, 1, ..., n. Since $\underline{V}^*(n+1) \in \mathcal{V}_n(\underline{\Sigma}^*(n))$, using lemma 3.3.4, (3.3.11) and (3.3.1), we have

$$E\{\underline{\tilde{e}}^*(n+1)\underline{y}'(n+1)\} = E\{\underline{\tilde{e}}^*(n+1)\underline{x}'(n+1)\underline{C}'(n+1)\}+E\{\underline{\tilde{e}}^*(n+1)\underline{\eta}'(n+1)\}$$

$$= -\underline{\Sigma}^*(n+1)\underline{C}'(n+1)+\underline{\tilde{V}}^*(n+1)\underline{Q}(n+1)$$

$$= -\underline{\Sigma}^*(n+1)\underline{C}'(n+1)+\underline{\Delta}^*(n+1)\underline{C}'(n+1)-\underline{\tilde{V}}^*(n+1)\underline{C}(n+1)\underline{\Sigma}^*(n)\underline{C}'(n+1)$$

$$= -\underline{\Sigma}^*(n+1)\underline{C}'(n+1)+\underline{\Sigma}^*(n+1)\underline{C}'(n+1) = \underline{0} \tag{3.3.27}$$

the corollary is proved by induction.

We are now in a position to prove (3.3.20). The results are stated as a theorem.

<u>Theorem 3.3.6</u>: Let $\{\underline{\tilde{e}}^*(k)\}_{k=0}^{\infty}$ be given by (3.3.21), where $\underline{\tilde{V}}^*(k) \in V_{k-1}(\underline{\Sigma}^*(k-1))$ with $\underline{\Sigma}^*(k-1)$ given by (3.3.11) and (3.3.12). Let $\{\underline{y}'(k)\}_{k=0}^{\infty}$ be given by (3.3.1) with $\underline{u}(k) = \underline{0}$, $k = 0, 1, \ldots$. Then for all $k = 0, 1, \ldots$, $\underline{\tilde{e}}^*(k)$ satisfies the projection equation; i.e.,

$$E\{\underline{\tilde{e}}^*(k)\underline{y}'(i)\} = \underline{0} \qquad i = 0, 1, \ldots, k \quad . \tag{3.3.28}$$

<u>Proof</u>: We shall use induction on k. By Corollary 3.3.5, (3.3.27) is true when k = 0.

Assume that (3.3.27) is true when k = 0, 1, ..., n. For i = 0, 1,..., n, we have from (3.3.21) and the induction hypothesis that

$$E\{\underline{\tilde{e}}^*(n+1)\underline{y}'(i)\} = (\underline{I}_n - \underline{\tilde{V}}^*(n+1)\underline{C}(n+1)\underline{A}(n)E\{\underline{\tilde{e}}^*(n)\underline{y}'(i)\} = \underline{0} \quad . \tag{3.3.29}$$

For i = n + 1, Corollary 3.3.7 gives

$$E\{\underline{\tilde{e}}^*(n + 1)\underline{y}'(n + 1)\} = \underline{0} \quad . \tag{3.3.30}$$

Combining (3.3.29) and (3.3.30), we have that (3.3.28) is true ;and the theorem is proved by induction.

In view of the Projection Theorem, this means that the estimate $\underline{w}^*(k)$ generated via $\mathcal{E}_T^2$, $T \in \mathcal{J}_{\tilde{V}^*}$, is the a.s. conditional mean of $\underline{x}(k), k=0,1,\ldots$. The results are also true when $\underline{u}(k)$, $k = 0, 1, \ldots$, is a nonzero but known control sequence, for one can always subtract off the deterministic contribution due to the nonzero known control sequence. The general situation where the $\underline{u}(k)$, $k = 0, 1, \ldots$, are generated by feedback of the observation sequence, shall be considered in detail in Chapter 5.

From the above discussions, we note that in the general case, there is more than one optimum classes of observers which will yield the same performance; only in the special case when $\underline{Q}(k) > \underline{0}$ or $\underline{C}(k+1)\underline{R}(k)\underline{C}'(k+1) > \underline{0}$, $k = 0, 1, \ldots$ (or both), there exists a unique optimum class of observers.

3.4 Minimal Order Optimum Observers for Stochastic Systems

Let $\tilde{\underline{V}}^*(k) \in \mathcal{U}_{k-1}(\underline{\Sigma}^*(k-1))$, $k = 0, 1, 2, \ldots$, specify an optimum class of observers. The class of minimum order optimum observers associated with $\{\tilde{\underline{V}}^*(k)\}_{k=0}^{\infty}$ is $\pi_{\tilde{V}^*}^p$, where p is the minimal order (or dimension). To find the number p amounts to finding the rank of the matrix $[\underline{I}_n - \tilde{\underline{V}}^*(k)\underline{C}(k)]$. We shall see that, depending on the observation noise we have that the minimal order optimum observers will have order which ranges from n - m to n.

Let us assume that the observations are partly deterministic, i.e.,

$$\underline{y}(k) = \begin{bmatrix} \underline{y}_1(k) \\ \cdots \\ \underline{y}_2(k) \end{bmatrix} = \begin{bmatrix} \underline{C}_1(k) \\ \cdots \\ \underline{C}_2(k) \end{bmatrix} \underline{x}(k) + \begin{bmatrix} \underline{n}_1(k) \\ \cdots \\ \underline{0} \end{bmatrix} \qquad (3.4.1)$$

where $\underline{y}_1(k) \in R^{m_1}$, $\underline{y}_2(k) \in R^{m_2}$. The vector $\underline{y}_2(k)$ is the noise-free component (Figure 3.2). This assumption has no loss of generality, for by appropriate transformation of the observation vector, all problems where the observation

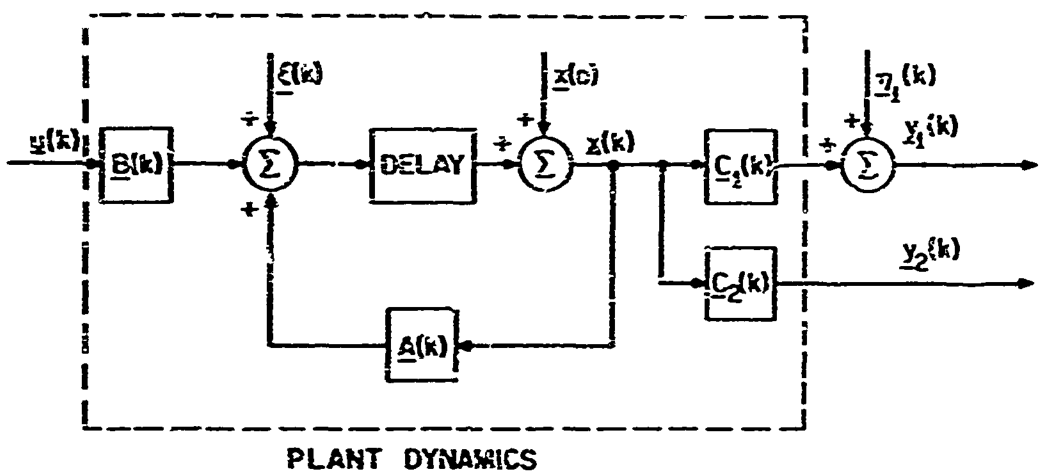


Fig. 3.2 THE STRUCTURE OF DISCRETE-TIME STOCHASTIC LINEAR SYSTEM S_2 WITH PARTIALLY NOISE-FREE OBSERVATION

noise is degenerate can be put into this form. We shall assume that

$$\begin{bmatrix} \underline{n}_1(k) \\ \cdots \\ \underline{0} \end{bmatrix}$$

has covariance matrix $\underline{Q}(k)$.

$$\underline{Q}(k) = \left[\begin{array}{c:c} \underline{Q}_1(k) & \underline{0} \\ \cdots & \cdots \\ \underline{0} & \underline{0}_{m_2 m_2} \end{array}\right] \quad ; \quad \underline{Q}_1(k) > \underline{0} \quad ; \quad \underline{Q}_1(k) \in M_{m_1 m_1} \quad . \tag{3.4.2}$$

Definition 3.4.1: A system with output $\underline{w}(k)$ is called compatible with respect to the noise-free observation $\underline{y}_2$ if for k = 0, 1, ...

$$\underline{C}_2(k)\underline{w}(k) = \underline{C}_2(k)\underline{x}(k) = \underline{y}_2(k) \quad \text{a.s.} \quad . \tag{3.4.3}$$

We shall also say that an observer $\mathcal{O}_T$, $T \in \mathcal{T}_V$, is compatible with respect to $\underline{y}_2$ if its associated estimator $\mathcal{E}_T$, $T \in \mathcal{T}_V$, is compatible with respect to $\underline{y}_2$.

Theorem 3.4.2: Let $\underline{\tilde{V}}^*(k) \in \mathcal{U}_{k-1}(\underline{\Sigma}^*(k-1))$, with $\underline{\Sigma}^*(k-1)$ given by (3.3.11) and (3.3.12); any $\mathcal{O}_T^2$, $T \in \mathcal{T}_{\tilde{V}^*}$ is compatible with respect to the noise-free observation $\underline{y}_2$.

Proof: Using Corollary 3.3.5, we have

$$\underline{\tilde{V}}^*(k)\underline{Q}(k) = \underline{\tilde{\Sigma}}^*(k)\underline{C}'(k) \quad ; \quad k = 0, 1, \ldots \quad . \tag{3.4.4}$$

Partition the matrix $\underline{\tilde{V}}^*(k)$ into

$$\underline{\tilde{V}}^*(k) = [\underline{\tilde{V}}_1^*(k) \,\vdots\, \underline{\tilde{V}}_2^*(k)] \quad ; \quad \underline{\tilde{V}}_1^*(k) \in M_{nm_1} \quad , \quad \underline{\tilde{V}}_2^*(k) \in M_{nm_2} \quad ;$$

$$k = 0, 1, \ldots \quad . \tag{3.4.5}$$

Equation (3.4.4) implies

$$\underline{V}_1^*(k)\underline{Q}_1(k) = \underline{\Sigma}^*(k)\underline{C}_1'(k) \quad ; \quad k = 0, 1, \ldots \tag{3.4.6}$$

$$\underline{0}_{nm_2} = \underline{\Sigma}^*(k)\underline{C}_2'(k) \quad ; \quad k = 0, 1, \ldots \tag{3.4.7}$$

where $\underline{\Sigma}^*(k)$ satisfies (3.3.11) and (3.3.12). The theorem follows from (3.4.7).

In the following, we shall consider the special case where the sequence of matrices $\{\underline{R}(k)\}_{k=0}^{\infty}$ are all positive definite. The general case can be treated in a similar approach. Since by assumption $\underline{R}(k) > \underline{0}$, $k = 0, 1, \ldots$, $\underline{\Sigma}^*(k)$ is given by equation (3.3.13), and $\underline{V}^*(k)$ given by (3.3.14) specifies the unique optimum class of observers.

<u>Lemma 3.4.2</u>: Let $\underline{R}(k) > \underline{0}$, $k = 0, 1, \ldots$. If the noise-free observation $\underline{y}_2(k) \in R^{m_2}$, $k = 0, 1, \ldots$, then $\underline{\Sigma}^*(k)$ given by (3.3.13) is of rank $n - m_2$, $k = 0, 1, \ldots$.

<u>Proof</u>: By compatibility (3.4.7) we have

$$\text{rank } \underline{\Sigma}^*(k) \le n - m_2 \qquad k = 0, 1, \ldots \quad . \tag{3.4.8}$$

From (3.4.6) and (3.3.14), we deduce that

$$\begin{aligned}\underline{C}_1(k)\underline{\Sigma}^*(k) &= \underline{Q}_1(k)\underline{V}_1^{*\prime}(k) \\ &= \underline{Q}_1(k)\underline{\Lambda}^*(k)\underline{C}_1(k)[\underline{\Delta}^*(k-1)-\underline{\Delta}^*(k-1)\underline{C}_2'(k) \\ &\qquad (\underline{C}_2(k)\underline{\Delta}^*(k-1)\underline{C}_2'(k))^{-1}\underline{C}_2(k)\underline{\Delta}^*(k-1)] \quad .\end{aligned} \tag{3.4.9}$$

where

$$\underline{\Lambda}^*(k) = \{\underline{Q}_1(k)+\underline{C}_1(k)[\underline{\Delta}^*(k-1)-\underline{\Delta}^*(k-1)\underline{C}_2'(k)$$

$$(\underline{C}_2(k)\underline{\Delta}^*(k-1)\underline{C}_2'(k))^{-1}\underline{C}_2(k)\underline{\Delta}^*(k-1)]\underline{C}_1'(k)\}^{-1} > 0 \quad ;$$

$$k = 0, 1, \ldots \quad . \quad (3.4.10)$$

(In deriving (3.4.9) from (3.3.14), a fair amount of matrix algebra is needed.) Let us define the matrix $\underline{\Gamma}^*(k - 1)$ by

$$\underline{\Gamma}^*(k-1) \triangleq \underline{\Delta}^*(k-1)\underline{C}_2'(k)(\underline{C}_2(k)\underline{\Delta}^*(k-1)\underline{C}_2'(k))^{-1}\underline{C}_2(k)\underline{\Delta}^*(k-1) \quad ;$$

$$k = 0, 1, \ldots \quad . \quad (3.4.11)$$

Now equation (3.4.9) can be written as

$$\underline{C}_1(k)\underline{\Sigma}^*(k) = \underline{Q}_1(k)\underline{\Lambda}^*(k)\underline{C}_1(k)[\underline{\Delta}^*(k-1)-\underline{\Gamma}^*(k-1)] \quad ; \quad k = 0, 1, \ldots \quad . \quad (3.4.12)$$

We note from (3.4.11) that if a vector $\underline{v} \in N(\underline{\Gamma}^*(k - 1))$, then it must be true that $\underline{\Delta}^*(k - 1)\underline{v} \in N(\underline{C}_2(k))$. Now suppose that the same vector $\underline{v} \in N(\underline{\Sigma}^*(k))$; then, from (3.4.9), we conclude that

$$\underline{Q}_1(k)\underline{\Lambda}^*(k)\underline{C}_1(k)\underline{\Delta}^*(k-1)\underline{v} = \underline{0} \implies \underline{\Delta}^*(k-1)\underline{v} \in N(\underline{C}_1(k))$$

$$k = 0, 1, \ldots \quad . \quad (3.4.13)$$

Therefore,

$$\underline{\Delta}^*(k - 1)\underline{v} \in N(\underline{C}(k)) \quad ; \quad k = 0, 1, \ldots \quad . \quad (3.4.14)$$

But from (3.3.11), we have

$$\underline{0} = \underline{\Delta}^*(k - 1)\underline{v} \quad ; \quad k = 0, 1, \ldots \quad . \quad (3.4.15)$$

Since $\underline{\Delta}^*(k - 1) > 0 \ \forall k$, we must have $\underline{v} = \underline{0}$. Thus,

$$N(\underline{\Gamma}^*(k - 1)) \cap N(\underline{\Sigma}^*(k)) = \{0\} \qquad k = 0, 1, \ldots \quad . \tag{3.4.16}$$

Clearly, the rank of $\underline{\Gamma}^*(k - 1)$ is m_2, and so $N(\underline{\Gamma}^*(k - 1))$ has dimension $n - m_2$ (3.4.16) implies that

$$\text{rank } (\underline{\Sigma}^*(k)) \geq n - m_2 \qquad k = 0, 1, \ldots \quad . \tag{3.4.17}$$

Equations (3.4.8) and (3.4.17) imply that rank $\underline{\Sigma}^*(k) = n - m_2$.

<u>Theorem 3.4.3</u>: Let $\underline{R}(k) > \underline{0}$, $k = 0, 1, \ldots$. If the noise-free observation $\underline{y}_2(k) \in R^{m_2}$, $k = 0, 1, \ldots$, then the class of minimal order observers is of order $n - m_2$.

<u>Proof</u>: From the remark made at the beginning of this section, one needs only to prove that the matrix $[\underline{I}_n - \underline{V}^*(k)\underline{C}(k)]$ has rank $n - m_2$, $k = 0, 1, \ldots$. (3.3.13) and (3.3.14) give us

$$[\underline{I}_n - \underline{V}^*(k)\underline{C}(k)]\underline{\Delta}^*(k - 1) = \underline{\Sigma}^*(k) \qquad , \qquad k = 0, 1, \ldots \quad . \tag{3.4.18}$$

By assumption, $\underline{R}(k) > \underline{0}$, $k = 0, 1, \ldots$; thus $\underline{\Delta}^*(k)$ has full rank for all k. By lemma 3.4.2, $\underline{\Sigma}^*(k)$ has rank $n - m_2$, $k = 0, 1, \ldots$; therefore, $[\underline{I}_n - \underline{V}^*(k)\underline{C}(k)]$ has rank $n - m_2$.

To end this section, we shall give <u>one</u> explicit minimal order optimum observer and its associated estimator for each case:

<u>Case 1</u>: $\underline{m_2 = 0}$

The class of minimal order optimum observer is of order n, and one explicit optimum estimator of minimal order can be constructed:

$$\mathcal{E}: \underline{w}^*(k+1) = (\underline{I}_n - \underline{V}^*(k+1)\underline{C}(k+1))\underline{A}(k)\underline{w}^*(k) + \underline{V}^*(k+1)\underline{y}(k+1) + \underline{B}(k)\underline{u}(k) \tag{3.4.19}$$

where $\{\underline{V}^*(k)\}_{k=0}^{\infty}$ is given by (3.3.11) to (3.3.13). <u>We notice that this is the Kalman estimator</u> [39]. (Figure 3.3.)

<u>Case 2</u>: $\underline{m_1 = 0}$

The class of minimal order optimum observer is of order n - m, i.e., there must exist $\underline{P}(k) \in M_{n(n-m)}$ and $\underline{T}(k) \in M_{(n-m)n}$, k = 0, 1, ... such that

$$[\underline{P}(k) \vdots \underline{V}^*(k)] \begin{bmatrix} \underline{T}(k) \\ \cdots \\ \underline{C}(k) \end{bmatrix} = \underline{I}_n \quad . \tag{3.4.20}$$

Since $\underline{T}(k) \in M_{(n-m)n}$ and $\underline{C}(k) \in M_{mn}$, (3.4.20) also implies that†

$$\underline{T}(k)\underline{V}^*(k) = \underline{0}_{m(n-m)} \quad ; \quad \underline{T}(k)\underline{P}(k) = \underline{I}_{n-m} \quad ; \quad \underline{C}(k)\underline{P}(k) = \underline{0}_{m(n-m)} \quad . \tag{3.4.21}$$

To specify <u>one</u> explicit minimal order optimum observer, let $\{\underline{P}^*(k)\}_{k=0}^{\infty}$ be any sequence of matrices such that

$$\underline{C}(k)\underline{P}^*(k) = \underline{0}_{m(n-m)} \qquad k = 0, 1, \ldots \quad . \tag{3.4.22}$$

Let $\{\underline{T}^*(k)\}_{k=0}^{\infty}$ be the solution of

$$\underline{T}^*(k)\underline{V}^*(k) = \underline{0}_{(n-m)m} \quad ; \quad \underline{T}^*(k)\underline{P}^*(k) = \underline{I}_{n-m} \quad ; \quad k = 0, 1, \ldots \quad . \tag{3.4.23}$$

The solution for (3.4.23) exists and is unique because we know a priori that (3.4.21) must have solutions (nonunique). The choice of $\{\underline{P}^*(k)\}_{k=0}^{\infty}$ is nonunique, and is usually chosen so as to simplify computation.

---

†Note that the condition $\underline{C}(k)\underline{V}^*(k) = \underline{I}_m$ is automatically implied by compatibility.

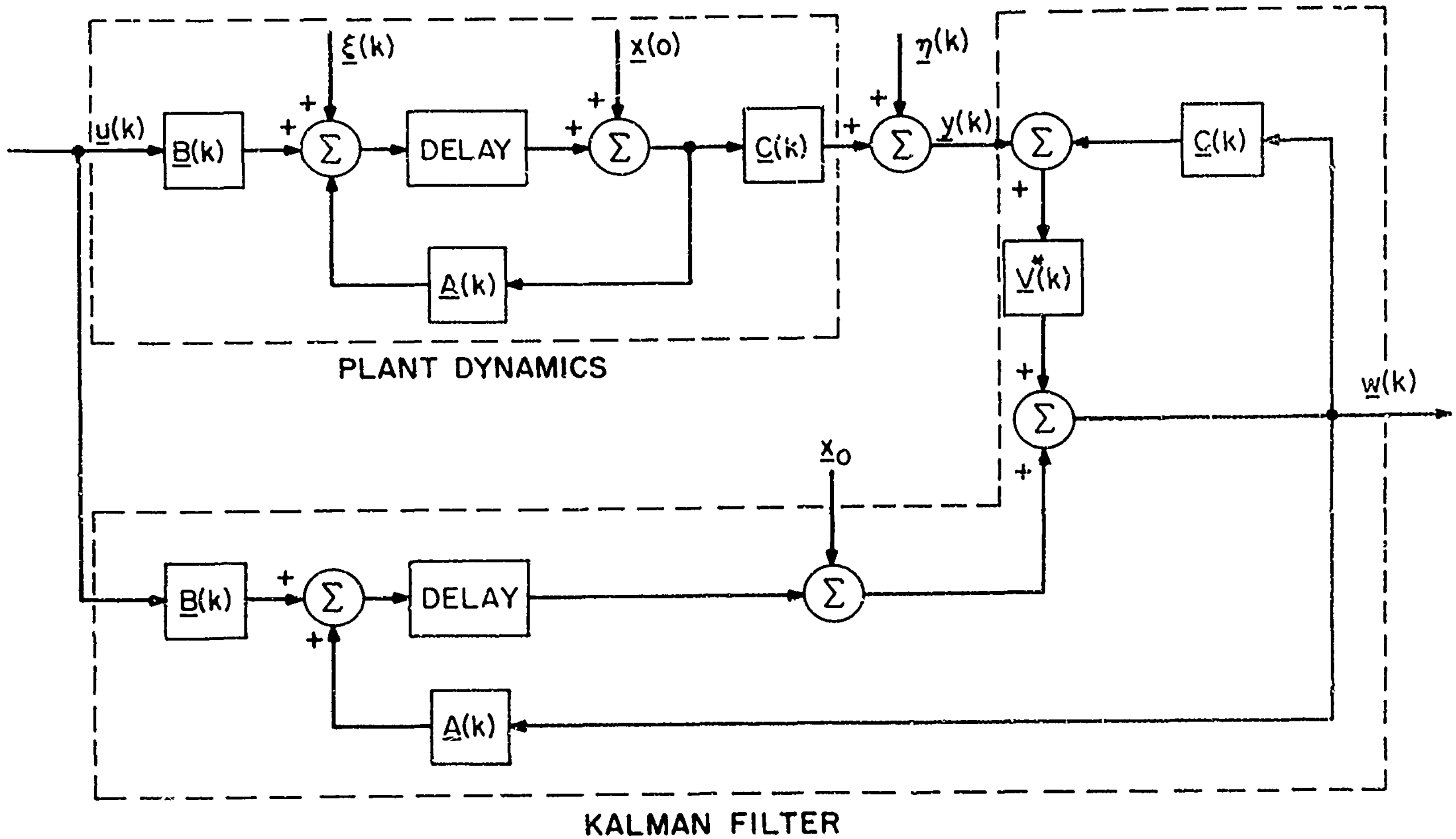


Fig. 3.3 STRUCTURE OF KALMAN FILTER. THE OBSERVATION NOISE IS NONDEGENERATE

Case 3: $0 < m_2 < m$

The class of minimal order optimum observer is of order $n - m_2$, i.e., there must exist $\underline{P}(k) \in M_{n(n-m_2)}$ and $\underline{T}(k) \in M_{(n-m_2)n}$, $k = 0, 1, \ldots$ such that

$$[\underline{P}(k) \vdots \underline{V}^*(k)] \begin{bmatrix} \underline{T}(k) \\ \cdots \\ \underline{C}(k) \end{bmatrix} = \underline{I}_n \tag{3.4.24}$$

and $\underline{T}(k)$ is of rank $n - m_2$. Choose $\underline{T}(k)$ such that

$$\underline{T}(k) = \begin{bmatrix} \underline{T}_1(k) \\ \cdots \\ \underline{T}_2(k) \end{bmatrix} \quad ; \quad \underline{T}_2(k) \in M_{(n-m)n} \tag{3.4.25}$$

and

$$\begin{bmatrix} \underline{T}_2(k) \\ \cdots \\ \underline{C}(k) \end{bmatrix}$$

is of rank n; thus $\underline{T}_1(k)$ must be given by

$$\underline{T}_1(k) = [\underline{K}_1(k) \vdots \underline{K}_2(k)] \begin{bmatrix} \underline{C}_1(k) \\ \cdots \\ \underline{C}_2(k) \end{bmatrix} \triangleq \underline{K}(k)\underline{C}(k) \tag{3.4.26}$$

where $\underline{K}_1(k) \in M_{m_1m_1}$, $\underline{K}_2(k) \in M_{m_1m_2}$. Partition $\underline{P}(k)$ into

$$\underline{P}(k) = [\underline{P}_1(k) \vdots \underline{P}_2(k)] \quad ; \quad \underline{P}_1(k) \in M_{nm_1} \quad , \quad \underline{P}_2(k) \in M_{n(n-m)} \quad . \tag{3.4.27}$$

Equations (3.4.24) to (3.4.27) imply also that

$$[\underline{P}_2(k) \;\vdots\; \underline{P}_1(k)\underline{K}(k) + \underline{V}^*(k)]\begin{bmatrix}\underline{T}_2(k)\\ \cdots \\ \underline{C}(k)\end{bmatrix} = \underline{I}_n \quad . \tag{3.4.28}$$

Since

$$\begin{bmatrix}\underline{T}_2(k)\\ \cdots \\ \underline{C}(k)\end{bmatrix} \varepsilon\, M_{nn} \quad ,$$

(3.4.28) also implies

$$\underline{T}_2(k)\underline{P}_2(k) = \underline{I}_{n-m} \quad ; \quad \underline{C}(k)\underline{P}_2(k) = \underline{0}_{m(n-m)} \quad ,$$

$$\underline{T}_2(k)\underline{P}_1(k)\underline{K}(k)+\underline{T}_2(k)\underline{V}^*(k) = \underline{0}_{(n-m)m} \quad ; \quad \underline{C}(k)\underline{P}_1(k)\underline{K}(k)+\underline{C}(k)\underline{V}^*(k) = \underline{I}_n \quad . \tag{3.4.29}$$

Partition $\underline{V}^*(k) = [\underline{V}_1^*(k) \;\vdots\; \underline{V}_2^*(k)]$, $\underline{V}_1^*(k) \,\varepsilon\, M_{nm_1}$, $\underline{V}_2^*(k) \,\varepsilon\, M_{nm_2}$; compatibility implies:

$$\underline{C}_2(k)\underline{P}_1(k) = \underline{0}_{m_2m_1} \quad ; \quad \underline{C}_2(k)\underline{V}_1^*(k) = \underline{0} \quad ; \quad \underline{C}_2(k)\underline{V}_2^*(k) = \underline{I}_{m_2} \quad . \tag{3.4.30}$$

Using (3.4.30), the last equation of (3.4.29) can be reduced to:

$$\underline{C}_1(k)\underline{P}_1(k)\underline{K}_2(k) + \underline{C}_1(k)\underline{V}_2^*(k) = \underline{0}_{m_1m_2} \tag{3.4.31}$$

$$\underline{C}_1(k)\underline{P}_1(k)\underline{K}_1(k) + \underline{C}_1(k)\underline{V}_1^*(k) = \underline{I}_{m_1} \quad . \tag{3.4.32}$$

Now to specify <u>one</u> explicit minimal order optimum observer let us choose $\{\underline{P}^*(k)\}_{k=0}^{\infty}$ such that

$$\underline{P}^*(k) = [\underline{P}_1^*(k) \;\vdots\; \underline{P}_2^*(k)] \quad ; \quad \underline{P}_1^*(k) \,\varepsilon\, M_{nm_1} \quad , \quad \underline{P}_2^*(k) \,\varepsilon\, M_{n(n-m)} \tag{3.4.33}$$

and $\underline{P}_1^*(k)$, $\underline{P}_2^*(k)$ satisfy:

$$\underline{C}(k)\underline{P}_2^*(k) = \underline{0}_{m(n-m)} \quad ; \quad \underline{C}_2(k)\underline{P}_1^*(k) = \underline{0}_{m_2 m_1} \quad . \tag{3.4.34}$$

Let $\underline{T}_2^*(k) \in M_{(n-m)n}$, $\underline{K}_1^*(k) \in M_{m_1 m_1}$, $\underline{K}_2^*(k) \in M_{m_1 m_2}$ be the solution of the following:

$$\begin{aligned}
\underline{C}_1(k)\underline{P}_1^*(k)\underline{K}_2^*(k) + \underline{C}_1(k)\underline{V}_2^*(k) &= \underline{0}_{m_1 m_2} \\
\underline{C}_1(k)\underline{P}_1^*(k)\underline{K}_1^*(k) + \underline{C}_1(k)\underline{V}_1^*(k) &= \underline{I}_{m_1} \\
\underline{T}_2^*(k)\underline{P}_1^*(k)\underline{K}_1^*(k) + \underline{T}_2^*(k)\underline{V}_1^*(k) &= \underline{0}_{(n-m)m_1} \quad ; \quad k = 0, 1, 2, \ldots \\
\underline{T}_2^*(k)\underline{P}_1^*(k)\underline{K}_2^*(k) + \underline{T}_2^*(k)\underline{V}_2^*(k) &= \underline{0}_{(n-m)m_2} \\
\underline{T}_2^*(k)\underline{P}_2^*(k) &= \underline{I}_{n-m} \quad .
\end{aligned} \tag{3.4.35}$$

Solution for (3.4.35) exists and is unique, since we know a priori that there are solutions for (3.4.29). The choice of $\{\underline{P}^*(k)\}_{k=0}^{\infty}$ is nonunique, and is usually chosen so as to simplify computation. One minimal order optimum estimator is $\mathcal{E}_{T^*}$, where

$$\underline{T}^*(k) = \begin{bmatrix} \underline{K}^*(k)\underline{C}(k) \\ \cdots \\ \underline{T}_2^*(k) \end{bmatrix} \tag{3.4.36}$$

and $\underline{K}^*(k)$, $\underline{T}_2^*(k)$ are given by the solution of (3.4.35). (Figure 3.4.)

3.5 Detectability and Observability of Linear Systems

Let us consider the totally noise free situation, i.e., $\underline{R}(k) = \underline{0}$ $\underline{Q}(k) = \underline{0}$, k = 0, 1, ... We shall discuss the notion of detectability and observability of the deterministic system $\mathcal{S}_1$ in terms of its structural properties.

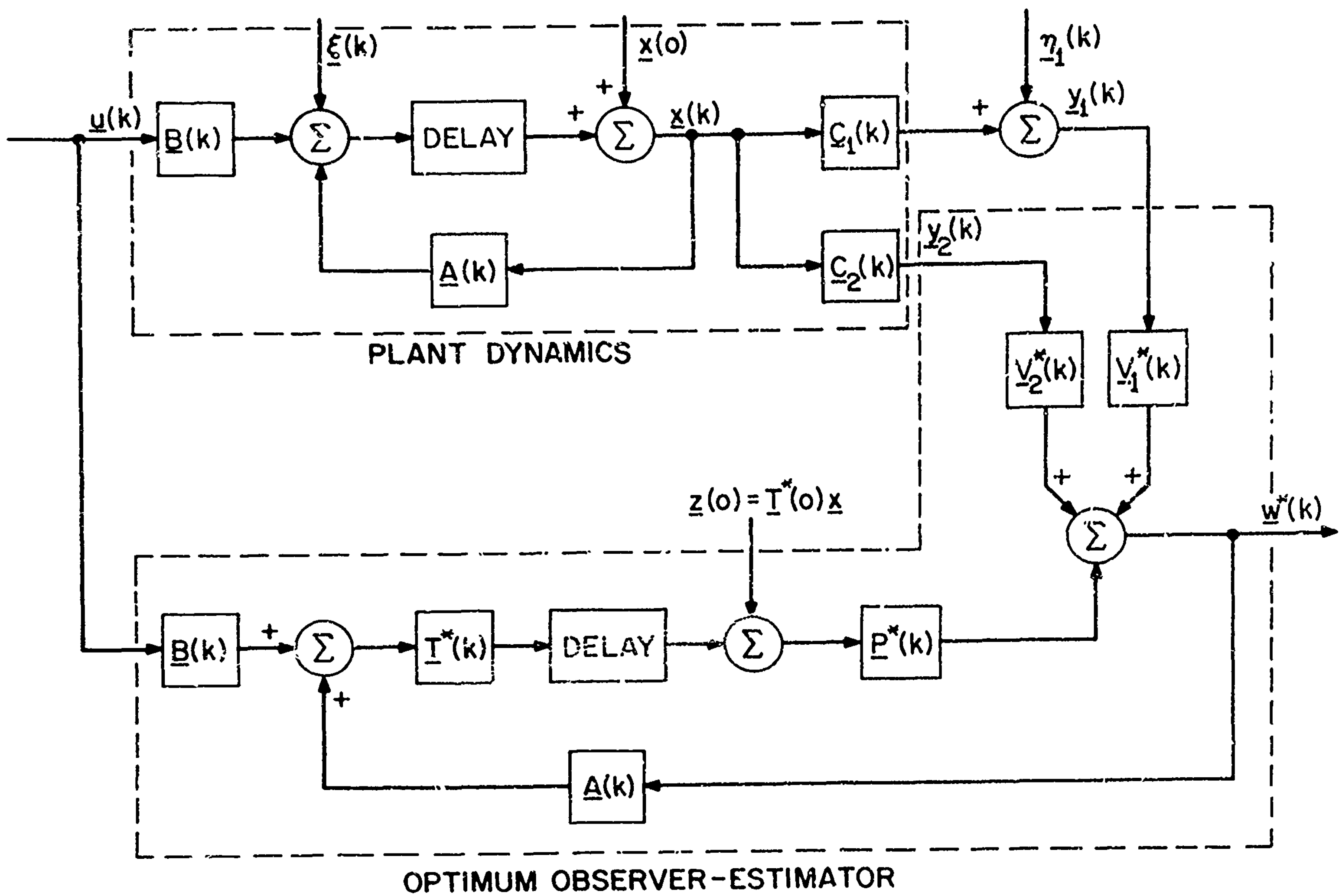


Fig. 3.4 STRUCTURE OF OPTIMUM OBSERVER-ESTIMATOR. THE OBSERVATION NOISE IS ASSUMED TO BE DEGENERATE IN GENERAL.

<u>Definition 3.5.1</u>: The system $\mathcal{S}_1$ is said to be detectable at k_o if for $\underline{x}(k_o) \in R^n$ arbitrary, there exists an estimator $\mathcal{E}_T^1$ described by

$$\mathcal{E}_T^1 \quad \begin{array}{l} \underline{z}(j+1) = \underline{T}(j+1)\underline{A}(j)\underline{P}(j)\underline{z}(j)+\underline{T}(j+1)\underline{A}(j)\underline{V}(j)\underline{y}(j)+\underline{T}(j+1)\underline{B}(j)\underline{u}(j) \\ \underline{w}(j) = \underline{P}(j)\underline{z}(j)+\underline{V}(j)\underline{y}(j) \quad ; \quad \underline{z}(k_o) \in S_{k_o} = \{\underline{T}(k_o)\underline{\alpha} | \underline{\alpha} \in R^n\} \end{array} \tag{3.5.1}$$

such that $\underline{w}(j) \to \underline{x}(j)$ as $j \to \infty$. The system $\mathcal{S}_1$ is said to be <u>detectable</u> if it is detectable at all $k_o = \ldots, -1, 0, 1, \ldots$.

<u>Definition 3.5.2</u>: The system $\mathcal{S}_1$ is said to be completely observable at k of index ν if for $\underline{x}(k) \in R^n$ arbitrary, we can deduce $\underline{x}(k)$ by observing $\underline{y}(k), \underline{y}(k+1), \ldots, \underline{y}(k+\nu-1)$.

One can easily show that an equivalent definition of observability is:[42]

<u>Definition 3.5.2'</u>: The system $\mathcal{S}_1$ is said to be completely observable at k_o of index ν if there exists $\nu < \infty$ such that

$$Q_{k_o,\nu} = [\underline{C}'(k_o) \vdots \underline{\phi}_A'(k_o,k_o)\underline{C}'(k_o+1) \vdots \ldots \vdots \underline{\phi}_A'(k_o+\nu-2,k_o)\underline{C}'(k_o+\nu-1)] \tag{3.5.2}$$

has rank n, where

$$\underline{\phi}_A(i,j) \triangleq \underline{A}(i)\underline{A}(i-1)\ldots\underline{A}(j) \quad ; \quad \underline{\phi}_A(i,i+1) \triangleq \underline{I}_n \quad ; \quad i \geq j \geq k_o \quad . \tag{3.5.3}$$

The system $\mathcal{S}_1$ is said to be <u>completely observable</u> if it is completely observable at all k_o with index ν_{k_o}, $k_o = \ldots, -1, 0, 1, \ldots$.

From the above definition, we cannot conclude <u>a priori</u> any relation between detectability and observability of the linear system. Intuitively, we may think complete observability implies detectability but at first sight, this implication is not obvious. In this section, we shall investigate the relation between observability and detectability.

Without loss of generality, we can set $k_o = 0$, and for simplicity write

$$\underline{Q}_k = \underline{Q}_{o,k} = [\underline{C}'(0) \vdots \underline{\Phi}_A'(0,0)\underline{C}'(1) \vdots \cdots \vdots \underline{\Phi}_A'(k-2,0)\underline{C}'(k-1)] \quad . \quad (3.5.4)$$

We shall assume that $\underline{A}(k)$ is invertible for all $k = \ldots, -1, 0, 1, \ldots$. Let us denote for $j \geq i$:

$$\underline{\Phi}_A(i,j) = \underline{\Phi}^{-1}(j,i) = \underline{A}^{-1}(j)\cdots\underline{A}^{-1}(j); \underline{\Phi}_A(i,i-1) \doteq \underline{I}_n \quad . \quad (3.5.5)$$

If we assume some a priori distribution on the initial condition of $\underline{x}(0)$, then we can make use of the results in Section 3.3 to obtain the equation for the error covariance, this is given by [see (3.3.11), (3.3.12)]

$$\begin{aligned} \underline{\Sigma}^*(k+1) &= \underline{\hat{\Sigma}}^*(k) - \underline{\hat{V}}^*(k+1)\underline{C}(k+1)\underline{\hat{\Sigma}}^*(k) \quad ; \quad k = 0, 1, \ldots \\ \underline{\Sigma}^*(0) &= \underline{\Sigma}_o - \underline{\Sigma}_o\underline{C}'(0)(\underline{C}(0)\underline{\Sigma}_o\underline{C}'(0))^{-1}\underline{C}(0)\underline{\Sigma}_o \end{aligned} \quad (3.5.6)$$

where

$$\underline{\hat{\Sigma}}^*(k) = \underline{A}(k)\underline{\Sigma}^*(k)\underline{A}'(k) \; ; \; k = 1, 2, \ldots \quad (3.5.7)$$

and

$$\underline{\hat{V}}^*(k) \in \mathcal{V}_{k-1}(\underline{\Sigma}^*(k-1)) = \{\underline{V} \in M_{nm} \, \underline{V}[\underline{C}(k)\underline{\hat{\Sigma}}^*(k-1)\underline{C}'(k)] = \underline{\hat{\Sigma}}^*(k-1)\underline{C}'(k)\} \quad ,$$

$$k = 0, 1, \ldots \quad .$$

Theorem 3.5.3: Let $\{\underline{\Sigma}^*(k)\}_0^\infty$ be a sequence satisfying (3.5.6) and (3.5.7) with $\underline{\hat{V}}^*(k) \in \mathcal{V}_{k-1}(\underline{\Sigma}^*(k-1))$. If $\underline{\Sigma}_o > \underline{0}$ (but arbitrary), then the null space of $\underline{\Sigma}^*(k)$ equals to the range space of $\underline{\Phi}_A'(0,k-1)\underline{Q}_k$, $k = 0, 1, \ldots$.

Proof: We shall use induction on k. For $k = 0$, we have from (3.5.6)

$$\underline{\Sigma}^*(0)\underline{Q}_o = \underline{\Sigma}^*(0)\underline{C}'(0) = \underline{\Sigma}_o\underline{C}'(0) - \underline{\Sigma}_o\underline{C}'(0) = \underline{0} \quad . \quad (3.5.8)$$

Thus we conclude that

$$\underline{N}(\underline{\Sigma}^*(0)) \supset R_a(\underline{Q}_o) \quad ; \quad \dim \underline{N}(\underline{\Sigma}^*(0)) \geq \dim R_a(\underline{Q}_o) = m \tag{3.5.9}$$

where $R_a(\cdot)$ denotes the range space. By assumption, $\underline{\Sigma}_o > \underline{0}$, thus from (3.5.6) we have

$$N(\underline{\Sigma}^*(0)) \cap N(\underline{\Sigma}_o\underline{C}'(0)(\underline{C}(0)\underline{\Sigma}_o\underline{C}'(0))^{-1}\underline{C}(0)\underline{\Sigma}_o) = \{\underline{0}\} \tag{3.5.10}$$

$\underline{C}(0)$ is of range m, therefore (3.5.10) implies

$$m \geq \dim N(\underline{\Sigma}^*(0)) \quad . \tag{3.5.11}$$

Equation (3.5.9) and (3.5.11) imply that

$$N(\underline{\Sigma}^*(0)) = R_a(\underline{Q}_o) = R(\underline{\psi}_A'(0,-1)\underline{Q}_o) \quad . \tag{3.5.12}$$

Let us assume that for k = i, we have

$$N(\underline{\Sigma}^*(i)) = R_a(\underline{\psi}_A'(0,i - 1)\underline{Q}_i) \quad . \tag{3.5.13}$$

From (3.5.4) and (3.5.5)

$$(\underline{A}^{-1}(i))'\underline{\psi}_A'(0,i - 1)\underline{Q}_i = \underline{\psi}_A'(0,i)\underline{Q}_i \quad . \tag{3.5.14}$$

Let $\underline{v} \in R_a(\underline{\psi}_A'(0,i)\underline{Q}_i)$, then there exists some $\underline{x} \in R^n$ such that

$$\underline{v} = \underline{\psi}_A'(0,i)\underline{Q}_i\underline{x} = (\underline{A}^{-1}(i))'\underline{\psi}_A'(0,i - 1)\underline{Q}_i\underline{x} \tag{3.5.15}$$

and so $\underline{A}'(i)\underline{v} \in R_a(\underline{\psi}_A'(0,i - 1)\underline{Q}_i)$, and, by the induction hypothesis, we also have $\underline{A}'(i)\underline{v} \in N(\underline{\Sigma}^*(i))$ or $\underline{v} \in N(\underline{\Delta}^*(i))$. By (3.5.6), we conclude $\underline{v} \in N(\underline{\Sigma}^*(i + 1))$. Also, by compatibility, the null space of $\underline{\Sigma}^*(i + 1)$ includes the range space of $\underline{C}'(i + 1)$. Combining the two, we have

$$\mathcal{N}(\underline{\Sigma}^*(i+1)) \supset R_a(\underline{\psi}_A'(0,i)\underline{Q}_{i+1}) \quad . \tag{3.5.16}$$

We also have the inequality

$$\dim \mathcal{N}(\underline{\Sigma}^*(i+1)) \geq \dim R_a(\underline{\psi}_A'(0,i)\underline{Q}_{i+1}) \quad . \tag{3.5.17}$$

Let $S_i = \{\underline{v} \in R^n | \underline{v} \in R_a(\underline{\Sigma}^*(i)) \cap \mathcal{N}(\underline{C}(i+1)\underline{A}(i))\}$. Since R^n is finite dimensional, from the induction hypothesis (3.5.13), we have $R_a(\underline{\Sigma}^*(i)) = \mathcal{N}(\underline{Q}_i'\underline{\psi}_A(0,i-1))$. Therefore any $\underline{v} \in S_i$ is described by

$$\underline{C}(i+1)\underline{A}(i)\underline{v} = \underline{0} \;\; ; \;\; \underline{Q}_i'\underline{\psi}_A(0,i-1)\underline{v} = \underline{0} \implies \underline{Q}_{i+1}'\underline{\psi}_A(0,i)\underline{A}(i)\underline{v} = \underline{0} \quad . \tag{3.5.18}$$

Since by assumption $\underline{A}(i)$ is nonsingular, equation (3.5.18) implies that

$$\dim S_i = \dim \mathcal{N}(\underline{Q}_{i+1}'\underline{\psi}_A(0,i)) = n - \dim R_a(\underline{\psi}_A'(0,i)\underline{Q}_{i+1}) \quad . \tag{3.5.19}$$

Let $\hat{S}_i^{-1}$ be the image of S_i through the transformation $\underline{\Sigma}^*(i)\underline{A}'(i)$; i.e., $\hat{S}_i^{-1} = \{\underline{w} \in R^n | \underline{\Sigma}^*(i)\underline{A}'(i)\underline{w} = \underline{v} \; ; \; \underline{v} \in S_i\}$. Let S_i^{-1} be the subspace which is equal to $\hat{S}_i^{-1}$ modulo the null space of $\underline{\Sigma}^*(i)\underline{A}'(i)$. S_i^{-1} has the same dimension as S_i, and so

$$\dim S_i^{-1} = n - \dim R_a(\underline{\psi}_A'(0,i)\underline{Q}_{i+1}) \quad . \tag{3.5.20}$$

Now let $\underline{w} \in S_i^{-1}$, $\underline{w} \neq 0$; then from the definition of S_i and S_i^{-1}, we have

$$\underline{C}(i+1)\underline{\Delta}^*(i)\underline{w} = \underline{0} \quad ; \quad \underline{\Delta}^*(i)\underline{w} \neq \underline{0} \quad . \tag{3.5.21}$$

From (3.5.6), we conclude

$$\underline{\Sigma}^*(i+1)\underline{w} = \underline{\Delta}^*(i)\underline{w} \neq \underline{0} \quad . \tag{3.5.22}$$

Therefore, the null space of $\underline{\Sigma}^*(i+1)$ and the space S_i^{-1} have only the zero element in common; thus

$$\dim N(\underline{\Sigma}^*(i+1)) \leq n - \dim S_i^{-1} = \dim R_a(\underline{\Phi}_A'(0,i)\underline{Q}_{i+1}) \quad . \qquad (3.5.23)$$

Equations (3.5.16), (3.5.17), and (3.5.23) imply

$$N(\underline{\Sigma}^*(i+1)) = R_a(\underline{\Phi}_A'(0,i)\underline{Q}_{i+1}) \quad . \qquad (3.5.24)$$

The theorem follows from induction.

A direct consequence of the above theorem is the following result :

Theorem 3.5.4: Let $\mathcal{S}_1$ be completely observable at time k of index ν_k; then there exists an optimum observer Θ_T', $T \in \mathcal{J}_V$, which will reconstruct the exact state, $\underline{x}(j)$, in at most ν_k steps (i.e., at time $k \leq j \leq k + \nu_k$). Thus if $\mathcal{S}_1$ is completely observable at time k of index ν_k, then $\mathcal{S}_1$ is detectable at time k; if S_1 is completely observable, then S_1 is detectable.

Theorem 3.5.4 **generalizes** Kalman's results [41] in deadbeat estimators; for this reason, we may refer to such an optimum observer Θ_T', $T \in \mathcal{J}_{\tilde{V}^*}$, as a deadbeat observer. Clearly there is more than one class of deadbeat ob- parameterized by $\{\tilde{\underline{V}}(k)\}_{k=0}^{\infty}$, $\tilde{\underline{V}}(k) \in \mathcal{V}_{k-1}(\underline{\Sigma}^*(k-1))$, $k = 1, 2, \ldots$. Among these, we shall find the simplest deadbeat observer.

Theorem 3.5.5: Let $\mathcal{S}_1$ be completely observable, the class of minimal order deadbeat observers is of order n-m .

Proof: Clearly, the class of minimal order deadbeat observer must be of order greater than or equal to n - m. To prove the theorem, we need to find a sequence $\{\tilde{\underline{V}}_1^*(k)\}_{k=0}^{\infty}$, $\tilde{\underline{V}}_1^*(k) \in \mathcal{V}_{k-1}(\underline{\Sigma}^*(k-1))$, $k = 1, \ldots$ such that the matrix $(\underline{I}_n - \tilde{\underline{V}}_1^*(k)\underline{C}(k))$ has rank less than or equal to n - m for all $k = 1, 2, \ldots$.

Construct

$$\underline{\tilde{V}}_1^*(k) = \underline{\Delta}^*(k-1)\underline{C}'(k)(\underline{C}(k)\underline{\Delta}^*(k-1)\underline{C}'(k))^{\#} + \underline{C}'(k)(\underline{C}(k)\underline{C}'(k))^{-1} \cdot$$
$$\{\underline{I}_m - \underline{C}(k)\underline{\Delta}^*(k-1)\underline{C}'(k)(\underline{C}(k)\underline{\Delta}^*(k-1)\underline{C}'(k))^{\#}\} \qquad k = 1, 2, \ldots \tag{3.5.25}$$

where $\underline{M}^{\#}$ denotes the pseudoinverse of a matrix $\underline{M}$. Using the properties of pseudoinverse (see Appendix A) we have

$$\begin{aligned}\underline{\tilde{V}}_1^*(k)\underline{C}(k)\underline{\Delta}^*(k-1)\underline{C}'(k) &= \underline{\Delta}^*(k-1)\underline{C}'(k)(\underline{C}(k)\underline{\Delta}^*(k-1)\underline{C}'(k))^{\#}\underline{C}(k)\underline{\Delta}^*(k-1)\underline{C}'(k) \\ &= \underline{\tilde{V}}^*(k)(\underline{C}(k)\underline{\Delta}^*(k-1)\underline{C}'(k))(\underline{C}(k)\underline{\Delta}^*(k-1)\underline{C}'(k))^{\#}\underline{C}(k)\underline{\Delta}^*(k-1)\underline{C}'(k) \\ &= \underline{\tilde{V}}^*(k)\underline{C}(k)\underline{\Delta}^*(k-1)\underline{C}'(k) = \underline{\Delta}^*(k-1)\underline{C}'(k); \quad k = 1, 2, \ldots\end{aligned} \tag{3.5.26}$$

Therefore $\underline{\tilde{V}}_1^*(k) \in \mathcal{V}_{k-1}(\underline{\Sigma}^*(k-1))$, $k = 1, 2, \ldots$. From (3.5.25), we deduce

$$\underline{C}(k)(\underline{I}_n - \underline{\tilde{V}}_1^*(k)\underline{C}(k)) = \underline{C}(k) - \underline{C}(k) = \underline{0} \quad . \tag{3.5.27}$$

Since $\underline{C}(k)$ is of rank m, (3.5.27) implies that

$$\text{rank } (\underline{I}_n - \underline{\tilde{V}}_1^*(k)\underline{C}(k)) \leq n - m \tag{3.5.28}$$

and the theorem follows.

Finally, we would like to derive a test for the detectability of linear systems. Using (3.3.7), we have easily the following:

Theorem 3.5.6: A system $\mathcal{S}_1$ is detectable if and only if there exists a uniformly bounded sequence $\{\underline{V}(k)\}_{k=0}^{\infty}$ such that

$$||\underline{\phi}_{\theta_V}(i,j)|| \leq \alpha_1 e^{-\alpha_2|i-j|} \quad ; \quad \alpha_1, \alpha_2 > 0 \tag{3.5.29}$$

where for $i \geq j$:

$$\underline{\theta}_k(\underline{V}(k+1)) \triangleq \underline{A}(k) - \underline{V}(k+1)\underline{C}(k+1)\underline{A}(k) \quad ; \quad \underline{\phi}_{\theta_V}(i,j) \triangleq \underline{\theta}_i(\underline{V}(i+1))\underline{\theta}_{i-1}(\underline{V}(i)) \; .$$

$$\ldots \underline{\theta}_j(\underline{V}(j+1)) \quad . \tag{3.5.30}$$

3.6 Asymptotic Behavior of Optimum Estimators

In this section, we shall investigate the asymptotic behavior of an optimum estimator $\mathcal{E}_T^2$, $T \in \mathcal{J}_{\tilde{V}^*}$, for the stochastic system $\mathcal{S}_2$. We shall say that the system $\mathcal{S}_2$ is detectable if its deterministic correspondence, $\mathcal{S}_1$, is detectable. The investigation is carried out by considering the minimal Riccati sequence $\{\underline{\Sigma}^*(k)\}_{k=0}^{\infty}$ which describes the evolution of the minimum error covariance.

First, let us assume that the initial time is k_o, and consider the behavior of $\underline{\Sigma}^*(k)$, as $k \to \infty$, where $\underline{\Sigma}^*(k)$ satisfies

$$\begin{aligned} \underline{\Sigma}^*(k+1) &= \underline{\Delta}^*(k) - \tilde{\underline{V}}^*(k+1)\underline{C}(k+1)\underline{\Delta}^*(k) \\ \underline{\Sigma}^*(k_o) &= \underline{\Sigma}_o - \underline{\Sigma}_o\underline{C}'(k_o)[\underline{C}(k_o)\underline{\Sigma}_o\underline{C}'(k_o) + \underline{Q}(k_o)]^{-1}\underline{C}(k_o)\underline{\Sigma}_o \end{aligned} \tag{3.6.1}$$

and $\tilde{\underline{V}}^*(k) \in V_{k-1}(\underline{\Sigma}^*(k-1))$, $k = k_o + 1, k_o + 2, \ldots$.

Theorem 3.6.1: The minimum covariance error $\underline{\Sigma}^*(k)$, will remain bounded for all $k = k_o + 1, k_o + 2, \ldots$ if and only if the system $\mathcal{S}_2$ is detectable.

Proof: If $\mathcal{S}_2$ is detectable, then from theorem 3.5.6, we note that there must exist a uniformly bounded sequence $\{\underline{V}(k)\}_{k=0}^{\infty}$ such that the resulting solution of (3.3.7) (with k_o replacing 0) will remain bounded for all $k = k_o, k_o + 1, \ldots$. Since $\{\underline{\Sigma}^*(k)\}_{k=0}^{\infty}$ satisfying (3.6.1) is the minimal sequence with respect to the solution set of (3.3.7), we conclude that $\underline{\Sigma}^*(k)$ must also be bounded for all $k = k_o, k_o + 1, \ldots$.

Conversely, if $\underline{\Sigma}^*(k)$ remains bounded for all $k = k_o, k_o + 1, \ldots$ then from (3.3.7), we have

$$||\underline{\phi}_{\underline{\theta}_{v*}}(i,j)|| \le \alpha_1 e^{-\alpha|i-j|} \qquad \alpha_1,\alpha_2 > 0 \tag{3.6.2}$$

and so $\mathcal{S}_2$ is detectable.

Next, we shall assume that the present time is k, and assume that the initial time $k_o \to -\infty$.

Let us rewrite (3.3.9) in a more suggestive form: (set $k_o' \triangleq k_o - 1$)

$$\underline{\Sigma}_v(k+1,k_o';\underline{\Sigma}_o) = (\underline{A}(k)-\underline{V}(k+1)\underline{C}(k+1)\underline{A}(k))\underline{\Sigma}_v(k+1,k_o';\underline{\Sigma}_o)(\underline{A}(k)-\underline{V}(k+1)\underline{C}(k+1))'$$

$$+(\underline{I}_n-\underline{V}(k+1)\underline{C}(k+1))\underline{R}(k)(\underline{I}_n-\underline{V}(k+1)\underline{C}(k+1))'+\underline{V}(k+1)\underline{Q}(k+1)\underline{V}'(k+1);\ k = k_o',k_o'+1,\ldots$$

$$\underline{\Sigma}_v(k_o',k_o';\underline{\Sigma}_o) = \underline{\Sigma}_o \ . \tag{3.6.3}$$

As we have noted, (3.6.3) is the same as (2.5.1) except for some changes in the symbol (3.3.14), (3.3.15). We shall still use the symbol as defined by (2.5.2) with the obvious change (3.3.14). As usual, we shall denote the minimal Riccati sequence with respect to the solution set of (3.6.3) by $\{\underline{\Sigma}^*(k,k_o';\Sigma_o)\}_{k=k_o'}^{\infty}$.

Lemma 3.6.2: There exists an unique bounded sequence $\{\underline{\Sigma}^*(k;\underline{0})\}_{k=-\infty}^{\infty}$ such that

$$\lim_{k_o'\to-\infty} \underline{\Sigma}^*(k,k_o';\underline{0}) = \underline{\Sigma}^*(k;\underline{0}) \qquad \text{for all } k \tag{3.6.4}$$

and $\underline{\Sigma}^*(k;\underline{0})$ satisfies

$$\underline{\Sigma}^*(k+1;\underline{0}) = \underline{\Phi}_k(\underline{\tilde{V}}^*(k+1),\underline{\Sigma}^*(k;\underline{0})) \ ; \ \underline{\tilde{V}}^*(k+1) \in \mathcal{V}_k(\underline{\Sigma}^*(k;\underline{0})) \tag{3.6.5}$$

if and only if the system $\mathcal{S}_2$ is detectable.

Proof: Let us denote

$$\underline{\theta}_k(\underline{V}) \triangleq \underline{A}(k) - \underline{V}\,\underline{C}(k+1)\underline{A}(k) \; ; \quad k = k_o', k_o' + 1, \ldots \quad . \qquad (3.6.6)$$

Using lemma 2.5.2 and Equation (3.6.3), we have the inequality

$$\underline{\theta}_{k-1}(\underline{\tilde{V}}^*(k,k_o'-1;\underline{0}))(\underline{\Sigma}^*(k-1,k_o'-1;\underline{0})-\underline{\Sigma}^*(k-1,k_o';\underline{0}))\underline{\theta}_{k-1}'(\underline{\tilde{V}}^*(k,k_o'-1;\underline{0}))^{\dagger}$$
$$\leq \underline{\Sigma}^*(k,k_o'-1,\underline{0})-\underline{\Sigma}^*(k,k_o';\underline{0}) \quad . \qquad (3.6.7)$$

Since $\underline{\Sigma}^*(k_o',k_o'-1;\underline{0}) \geq \underline{0}$ and $\underline{\Sigma}^*(k_o',k_o';\underline{0}) = \underline{0}$, (3.6.7) implies that

$$\underline{\Sigma}^*(k,k_o-1;\underline{0}) \geq \underline{\Sigma}^*(k,k_o;\underline{0}) \qquad \text{for all } k \geq k_o \quad . \qquad (3.6.8)$$

If $\mathcal{S}_2$ is detectable, then by using theorem 3.6.1, $\underline{\Sigma}^*(k,k_o;\underline{0})$ will be bounded for a fixed k and all $k_o \leq k$. By the monotone convergence theorem of nonnegative operators [32], we conclude that there exists an unique $\underline{\Sigma}^*(k;\underline{0})$ such that

$$\lim_{k_o'\to-\infty} \underline{\Sigma}^*(k,k_o';\underline{0}) = \underline{\Sigma}^*(k;\underline{0}) \geq \underline{\Sigma}^*(k,k_o;\underline{0}) \quad ; \quad k_o > -\infty \quad . \qquad (3.6.9)$$

Let us define

$$\underline{\hat{\Sigma}}^*(k+1;\underline{0}) = \underline{\psi}_k(\underline{\tilde{V}}^*(k+1),\underline{\Sigma}^*(k;\underline{0})) \quad ; \quad \underline{\tilde{V}}^*(k+1) \in \mathcal{U}_k(\underline{\Sigma}^*(k;\underline{0})) \quad . \qquad (3.6.10)$$

By lemma 2.5.2 and (3.6.8) we have for all $k_o' > -\infty$:

$$\underline{\hat{\Sigma}}^*(k+1;\underline{0})-\underline{\Sigma}^*(k+1;\underline{0}) \leq \underline{\psi}_k(\underline{\tilde{V}}^*(k+1),\underline{\Sigma}^*(k;\underline{0}))-\underline{\psi}_k(\underline{\tilde{V}}^*(k+1,k_o';\underline{0}),\underline{\Sigma}^*(k,k_o';\underline{0}))$$
$$\leq \underline{\psi}_k(\underline{\tilde{V}}^*(k+1,k_o';\underline{0}),\underline{\Sigma}^*(k;\underline{0}))-\underline{\psi}_k(\underline{\tilde{V}}^*(k+1,k_o';\underline{0}),\underline{\Sigma}^*(k,k_o';\underline{0})) \quad . \qquad (3.6.11)$$

$^{\dagger}\underline{\tilde{V}}^*(k,k_o'-i;\underline{0}) \in \mathcal{U}_{k-1}(\underline{\Sigma}^*(k-1,k_o'-i;\underline{0})) \quad , \quad i = 0, 1 \quad .$

Since $\underline{\Sigma}^*(k,k_o';\underline{0}) \to \underline{\Sigma}^*(k;\underline{0})$ as $k_o' \to -\infty$, (3.6.11) implies

$$\hat{\underline{\Sigma}}^*(k+1;\underline{0}) \le \underline{\Sigma}^*(k+1;\underline{0}) \quad . \tag{3.6.12}$$

From (2.5.12) (3.6.9) and lemma 2.5.2, we have for all k_o':

$$\hat{\underline{\Sigma}}^*(k+1;\underline{0}) \ge \underline{\psi}_k(\tilde{\underline{V}}^*(k+1),\underline{\Sigma}^*(k,k_o';\underline{0})) \ge \underline{\psi}_k(\tilde{\underline{V}}^*(k+1,k_o';\underline{0}),\underline{\Sigma}^*(k,k_o';\underline{0})) = \underline{\Sigma}^*(k+1,k_o';\underline{0}) \tag{3.6.13}$$

and so taking $k_o' \to -\infty$:

$$\hat{\underline{\Sigma}}^*(k+1;\underline{0}) \ge \underline{\Sigma}^*(k+1;\underline{0}) \quad . \tag{3.6.14}$$

Combining (3.6.12) and (3.6.14) we obtain (3.6.5).

Conversely if (3.6.4) and (3.6.5) are true, then it must be true that

$$||\phi_{\theta_{\tilde{V}^*}}(i,j)|| \le \alpha_1 e^{-\alpha_2|i-j|} \qquad \alpha_1,\alpha_2 > 0 \tag{3.6.15}$$

and so $\mathcal{S}_2$ is detectable by theorem 3.5.6.

<u>Theorem 3.6.3</u>: There exists a unique sequence $\{\underline{\Sigma}^*(k)\}_{k=-\infty}^{\infty}$ such that

$$\lim_{k_o' \to -\infty} \underline{\Sigma}^*(k,k_o;\underline{\Sigma}_o) = \underline{\Sigma}^*(k) \qquad \text{for all } k \tag{3.6.16}$$

with $\underline{\Sigma}_o \ge \underline{0}$; and $\underline{\Sigma}^*(k)$ satisfies

$$\underline{\Sigma}^*(k+1) = \underline{\psi}_k(\tilde{\underline{V}}^*(k+1),\underline{\Sigma}^*(k)) = \underline{\Delta}^*(k)-\tilde{\underline{V}}^*(k+1)\underline{C}(k+1)\underline{\Delta}^*(k) \quad ; \quad \tilde{\underline{V}}^*(k+1) \in \mathcal{V}_k(\underline{\Sigma}^*(k))$$

if and only if the system $\mathcal{S}_2$ is detectable.

<u>Proof</u>: Using lemma 2.5.2, we have

$$\underline{\theta}_k(\tilde{\underline{V}}^*(k,k_o';\underline{\Sigma}_o))(\underline{\Sigma}^*(k,k_o';\underline{\Sigma}_o)-\underline{\Sigma}^*(k,k_o';\underline{0}))\underline{\theta}_k'(\tilde{\underline{V}}^*(k,k_o';\underline{\Sigma}_o)) \le \underline{\Sigma}^*(k+1,k_o';\underline{\Sigma}_o)-$$
$$\underline{\Sigma}^*(k+1,k_o';\underline{0}) \le \underline{\theta}_k(\tilde{\underline{V}}^*(k,k_o';\underline{0}))(\underline{\Sigma}^*(k,k_o';\underline{\Sigma}_o)-\underline{\Sigma}^*(k,k_o';\underline{0}))\underline{\theta}_k'(\tilde{\underline{V}}^*(k,k_o';\underline{0})) \quad . \tag{3.6.18}$$

Since $\underline{\Sigma}^*(k_o',k_o';\underline{\Sigma}_o) = \underline{\Sigma}_o \geq \underline{0}$, then (3.6.18) implies that

$$\underline{0} \leq \underline{\Sigma}^*(k+1,k_o';\underline{\Sigma}_o) - \underline{\Sigma}^*(k+1,k_o';\underline{0}) \leq \underline{\phi}_\theta(k,k_o')\underline{\Sigma}_o\underline{\phi}_\theta'(k,k_o')$$

where

$$\underline{\phi}_\theta(k,k_o') \triangleq \underline{\theta}_k(\underline{\tilde{V}}^*(k,k_o';\underline{0}))\underline{\theta}_{k-1}(\underline{\tilde{V}}^*(k-1,k_o';\underline{0})\ldots\underline{\theta}_{k_o'}(\underline{\tilde{V}}^*(k_o',k_o';\underline{0}) \quad . \tag{3.6.20}$$

If the system S_2 is detectable, then by theorem 3.6.1, we must have

$$||\underline{\phi}_\theta(k,k_o')|| \leq \alpha_1 e^{-\alpha_2|k-k_o'|} \qquad \alpha_1,\alpha_2 > 0 \tag{3.6.21}$$

and so using lemma 3.6.2 and equations (3.6.19), (3.6.21), we have

$$\lim_{k_o'\to-\infty} \underline{\Sigma}^*(k,k_o';\underline{\Sigma}_o) = \lim_{k_o'\to-\infty} \underline{\Sigma}^*(k,k_o';\underline{0}) = \underline{\Sigma}^*(k;\underline{0}) \triangleq \underline{\Sigma}^*(k) \quad . \tag{3.6.22}$$

Equation (3.6.17) follows from (3.6.5).

The proof in the reverse direction is the same as in proving lemma 3.6.2.

Finally, we shall consider the <u>time invariant</u> case where $\underline{A}$, $\underline{C}$, $\underline{Q}$, $\underline{R}$ are constant and bounded matrices. In this case

$$\underline{\Sigma}^*(k,k_o';\underline{\Sigma}_o) = \underline{\Sigma}^*(k-k_o',0;\underline{\Sigma}_o) \tag{3.6.23}$$

thus taking $k_o' = 0$ is the same as considering $k \to \infty$. We shall only consider $k_o' = 0$, and $k \to \infty$.

<u>Theorem 3.6.4</u>: There exists a bounded $\underline{\Sigma}^*$ such that

$$\lim_{k\to\infty} \underline{\Sigma}^*(k,0;\underline{\Sigma}_o) = \underline{\Sigma}^* \tag{3.6.24}$$

and $\underline{\Sigma}^*$ satisfies the algebraic equation

$$\underline{\Sigma}^* = \underline{\Delta}^* - \underline{\tilde{V}}^*\underline{C}\,\underline{\Delta}^* \quad ; \quad \underline{\tilde{V}}^* \in \mathcal{V}(\underline{\Sigma}^*) \triangleq \{\underline{V} \in M_{nm} | \underline{V}(\underline{C}\,\underline{\Delta}^*\underline{C}' + \underline{Q}) = \underline{\Delta}\,\underline{C}'\} \tag{3.6.25}$$

$$\underline{\Delta}^* = \underline{A}\,\underline{\Sigma}^*\underline{A}' + \underline{Q} \tag{3.6.26}$$

if and only if the system S_2 is detectable.

Proof: In the time invariant case

$$\underline{\psi}_k(\underline{V},\underline{\Sigma}) = \underline{\psi}_{k+1}(\underline{V},\underline{\Sigma}) \triangleq \underline{\psi}(\underline{V},\underline{\Sigma}) \quad ; \quad \underline{\theta}_k(\underline{V}) = \underline{\theta}_{k+1}(\underline{V}) \triangleq \underline{\theta}(\underline{V}) \quad . \tag{3.6.27}$$

Using lemma 2.5.2, we have

$$\underline{\Sigma}^*(k+1,0;\underline{0}) - \underline{\Sigma}^*(k,0;\underline{0}) \geq \underline{\theta}(\underline{\tilde{V}}^*(k,0;\underline{0}))(\underline{\Sigma}^*(k,0;\underline{0}) - \underline{\Sigma}^*(k-1,0;\underline{0}))$$
$$\underline{\theta}'(\underline{\tilde{V}}^*(k,0;\underline{0})) \quad . \tag{3.6.28}$$

Since $\underline{\Sigma}^*(1,0;\underline{0}) \geq \underline{0}$, (3.6.28) implies that

$$\underline{\Sigma}^*(k+1,0;\underline{0}) \geq \underline{\Sigma}^*(k,0;\underline{0}) \qquad k = 0,\ 1,\ 2,\ \ldots \quad . \tag{3.6.29}$$

By theorem 3.6.1, $\underline{\Sigma}^*(k,0;\underline{0})$ will remain bounded if and only if S_2 is detectable, and so by (3.6.29), one concludes that there exists $\underline{\Sigma}^*$ such that

$$\lim_{k\to\infty} \underline{\Sigma}^*(k,0;\underline{0}) = \underline{\Sigma}^* \quad . \tag{3.6.30}$$

Using (3.6.23) and lemma 3.6.2, $\underline{\Sigma}^*$ satisfies the algebraic equation

$$\underline{\Sigma}^* = \underline{\psi}(\underline{\tilde{V}}^*,\underline{\Sigma}^*) = \underline{\Delta}^* - \underline{\tilde{V}}^*\underline{C}\,\underline{\Delta}^* \quad ; \quad \underline{\tilde{V}}^* \in \mathcal{V}(\underline{\Sigma}^*) \tag{3.6.31}$$

and $\underline{\Sigma}^*$ is given by (3.6.26) if and only if S_2 is detectable. Using theorem 3.6.3 we have the desired results.

Theorem 3.6.5: If S_2 is detectable, there exists only one nonengative definite matrix $\underline{\Sigma}^*$ which satisfies (3.6.25), (3.6.26).

<u>Proof</u>: Let us define

$$\underline{\Sigma}^{o} = \underline{0} \quad ; \quad \underline{V}^{o} \in V(\underline{0}) \tag{3.6.32}$$

and

$$\underline{\Sigma}^{i} = \psi(\underline{V}^{i-1}, \underline{\Sigma}^{i-1}) \quad ; \quad \underline{V}^{i} \in \mathcal{V}(\underline{\Sigma}^{i}) \quad . \tag{3.6.33}$$

By (3.6.30), such a constructed sequence of $\underline{\Sigma}^{i}$ will converge to $\underline{\Sigma}^{*}$, which satisfies (3.6.25), (3.6.26).

Let $\underline{\tilde{\Sigma}} \geq \underline{0}$, and $\underline{\tilde{\Sigma}}$ satisfies (3.6.25), (3.6.26); i.e.,

$$\underline{\tilde{\Sigma}} = \underline{\psi}(\underline{\tilde{V}}, \underline{\tilde{\Sigma}}) \quad ; \quad \underline{\tilde{V}} \in \mathcal{V}(\underline{\tilde{\Sigma}}) \quad . \tag{3.6.24}$$

By lemma 2.5.2, we have

$$\underline{\theta}(\underline{\tilde{V}})(\underline{\tilde{\Sigma}} - \underline{\Sigma}^{i-1})\underline{\theta}'(\underline{\tilde{V}}) \leq \underline{\tilde{\Sigma}} - \underline{\Sigma}^{i} \leq \underline{\theta}(\underline{V}^{i})(\underline{\tilde{\Sigma}} - \underline{\Sigma}^{i-1})\underline{\theta}'(\underline{V}^{i}) \quad . \tag{3.6.35}$$

By construction, $\underline{\Sigma}^{o} = \underline{0}$, thus

$$\underline{0} \leq \underline{\tilde{\Sigma}} - \underline{\Sigma}^{i} \leq \underline{\phi}_{\theta}(i,0)\underline{\tilde{\Sigma}}\, \underline{\phi}'_{\theta}(i,0) \tag{3.6.36}$$

where

$$\underline{\phi}_{\theta}(i,0) \triangleq \underline{\theta}(\underline{V}^{i})\underline{\theta}(\underline{V}^{i-1}) \ldots \underline{\theta}(\underline{V}^{o}) \quad . \tag{3.6.37}$$

If S_2 is detectable, then

$$||\underline{\phi}_{\theta}(i,j)|| \leq \alpha_1 e^{-\alpha_2|i-j|} \qquad \alpha_1, \alpha_2 > 0 \tag{3.6.38}$$

thus we have

$$\underline{\Sigma}^{i} \rightarrow \underline{\tilde{\Sigma}} \quad \text{as } i \rightarrow \infty \tag{3.6.39}$$

and uniqueness follows.

3.7 General Discussion

In this chapter, we have obtained an optimum unbiased estimator for the stochastic system S_z, where the observation noise may be degenerate ($Q(k) \geq 0$) or singular ($C(k) = 0$). In essence, the optimum estimator is specified by the relations:

$$E_T^2 : \quad \begin{aligned} z^*(k+1) &= T(k+1)A(k)P(k)z^*(k) + T(k+1)A(k)\tilde{V}^*(k)y(k) + T(k+1)B(k)u(k) \\ x^*(k) &= P(k)z^*(k) + \tilde{V}^*(k)y(k) \quad ; \quad z(0) = T(0)x_o \end{aligned} \tag{3.7.1}$$

where $x^*(k)$ is the optimal estimates of $x(k)$. (See Figure 3.4.) The matrices $P(k)$, $T(k)$ satisfy

$$P(k)T(k) + \tilde{V}^*(k)C(k) = I_n \tag{3.7.2}$$

and $\tilde{V}^*(k)$ is given by (see theorem 3.3.1)

$$\begin{aligned} &\tilde{V}^*(k) \in L_{k-1}(\Sigma^*(k-1)) \quad , \quad k = 1, 2, \ldots \quad ; \quad \tilde{V}^*(0) = \Sigma_o C'(0)[C(0)\Sigma_o C'(0)+Q(0)]^{-1} \\ &\Sigma^*(k+1) = \Delta^*(k) - \tilde{V}^*(k+1)C(k+1)\Delta^*(k) \quad ; \\ &\qquad \Sigma(0) = \Sigma_o - \Sigma_o C'(0)[C(0)\Sigma_o C'(0)+Q(0)]^{-1}C(0)\Sigma_o \\ &\Delta^*(k) \triangleq A(k)\Sigma^*(k)A'(k)+R(k) \quad . \end{aligned} \tag{3.7.3}$$

Note that $\{\tilde{V}^*(k)\}_{k=0}^{\infty}$ can be precomputed when the structure of S_z and the statistical law of the uncertainties are known. In general, $\{\tilde{V}^*(k)\}_{k=0}^{\infty}$ may not be unique. In the special case when $Q(k) > 0$ or $C(k+1)R(k)C'(k+1) > 0$ we have uniqueness in $\{\tilde{V}^*(k)\}_{k=0}^{\infty}$ (see theorem 3.3.1).

Once $\{\tilde{V}^*(k)\}_{k=0}^{\infty}$ is found, we can choose different $\{P(k)\}_{k=0}^{\infty}$ and $\{T(k)\}_{k=0}^{\infty}$ such that (3.7.2) is satisfied, and so one can construct different

optimum estimators $\mathcal{E}_T^2$, where the dimension of the observer state vector $\underline{z}(k)$ are different depending on the choice of $\{\underline{T}(k)\}_{k=0}^{\infty}$, $\underline{T}(k) \in \mathcal{J}_{\tilde{\underline{\nu}}^*(k)}$. It has been shown that in the special case when $\underline{R}(k) > \underline{0}$, the minimal order optimum observer is of dimension $n - m_2$ where m_2 is the number of noise-free channels available. Though the proof is given for the special case, it is conjectured that the results will be true in the general case when $\underline{R}(k) \geq \underline{0}$ or even $\underline{R}(k) = \underline{0}$.

Finally, the asymptotic behavior of $\underline{\Sigma}^*(k)$ given by (3.7.3) was considered in great detail. Necessary and sufficient condition were derived for $\underline{\Sigma}^*(k)$ to be uniformly bounded and existence of its steady behavior.

In the following, we shall discuss some of the relevant points in the development of this chapter.

(A) Discussion of Approaches

Different approaches are available to filtering problems. The Projection approach was used by Kalman to first obtain the Kalman filter. The starting point of this approach is the Projection Theorem (Theorem 3.3.2). There is also the Baysian approach [43] where one computes the conditional expectation of the state, $\underline{x}(k)$. Also, a max-likelihood approach [44] is available to filtering problems. Then, there is the approach of unbiased minimum error covariance estimates [10], and of weighted least square estimates [43]. In the linear-Gaussian case all these approaches will yield the same solution (see also section 3.3). It is hard to argue which of the above approachs to the problem is more fundamental than the other, for this highly depends on one's philosophical viewpoint to the problem. One may argue that the Baysian approach is the most fundamental approach. This is true to the extent where one can justify

the knowledge on a priori distribution of all the underlying random vectors.

The approach used in this chapter seems to be a new approach to the problem where one starts from deterministic consideration This is true in some sense. If the knowledge on the a priori distribution of the state $\underline{x}(0)$ is correct, then the approach is equivalent to that of unbiased minimum error covariance. To verify this statement let us consider the stochastic system $\mathcal{S}_2$ (with $\underline{u}(k) \equiv \underline{0}$). We look for an unbiased estimator which is nonanticipative. In general, such an estimator is described by [45]:

$$\mathcal{E}: \begin{cases} \underline{z}(k+1) = \underline{F}(k)\underline{z}(k) + \underline{G}(k)\underline{y}(k) \quad ; \quad \underline{z}(k) \in R^s \\ \underline{w}(k) = \underline{P}(k)\underline{z}(k) + \underline{V}(k)\underline{y}(k) \end{cases} \quad . \tag{3.7.4}$$

The initial condition of $\underline{z}(0)$ is some linear transformation of $\underline{x}_o$; i.e.,

$$\underline{z}(0) = \underline{T}(0)\underline{x}_o \tag{3.7.5}$$

and for all $k \geq 0$, we want $E\{\underline{w}(k)\} = E\{\underline{x}(k)\}$. With this restriction we have

$$(\underline{P}(0)\underline{T}(0) + \underline{V}(0)\underline{C}(0)\underline{x}_o = \underline{x}_o \quad . \tag{3.7.6}$$

We would like to construct the estimator completely independent of the mean of $\underline{x}(0)$, then (3.7.6) implies

$$\underline{P}(0)\underline{T}(0) + \underline{V}(0)\underline{C}(0) = \underline{I}_n \tag{3.7.7}$$

and so we must have $s \geq n - m$ and $\underline{T}(0) \in \Omega(\underline{C}(0);m,s,n)$. For $k > 0$, the unbiased restriction gives

$$\underline{P}(k)[\underline{\phi}_F(k-1,0)\underline{T}(0)\underline{x}_o + \sum_{i=0}^{k-1} \underline{\phi}_F(k-1,i+1)\underline{G}(i)\underline{C}(i)\underline{\phi}_A(i-1,0)\underline{x}_o]$$

$$+ \underline{V}(k)\underline{C}(k)\underline{\phi}_A(k-1,0)\underline{x}_o = \underline{\phi}_A(k-1,0)\underline{x}_o \quad . \qquad (3.7.8)$$

If $\underline{A}(k)$ are invertible for all k, and the structure of the estimator is independent of $\underline{x}_o$, then (3.7.8) implies that

$$\underline{P}(k)[\underline{\phi}_F(k-1,0)\underline{T}(0)\underline{\phi}_A(0,k-1) + \sum_{i=0}^{k-1} \underline{\phi}_F(k-1,i+1)\underline{G}(i)\underline{C}(i)\underline{\phi}_A(i,k-1)]$$

$$+ \underline{V}(k)\underline{C}(k) = \underline{I}_n \quad . \qquad (3.7.9)$$

Define

$$\underline{T}(k) = \underline{\phi}_F(k-1,0)\underline{T}(0)\underline{\phi}_A(0,k-1) + \sum_{i=0}^{k-1} \underline{\phi}_F(k-1,i+1)\underline{G}(i)\underline{C}(i)\underline{\phi}_A(i,k-1) \quad . \qquad (3.7.10)$$

Then $\underline{T}(k) \in \mathcal{J}(\underline{C}(k);m,s,n)$ and $\underline{T}(k)$ satisfies:

$$\underline{T}(k+1) = \underline{F}(k)\underline{T}(k)\underline{A}^{-1}(k) + \underline{D}(k)\underline{C}(k)\underline{A}^{-1}(k) \quad . \qquad (3.7.11)$$

Such an estimator can be realized by picking

$$\underline{F}(k) = \underline{T}(k+1)\underline{A}(k)\underline{P}(k) \quad ; \quad \underline{D}(k) = \underline{T}(k+1)\underline{A}(k)\underline{V}(k) \quad . \qquad (3.7.12)$$

Comparing with theorem 3.2.3, we see that all unbiased, nonanticipative estimators can be realized by an observer Θ^2_Γ, $T \in \mathcal{J}_V$, and its associated estimator $\mathcal{E}^2_T$. Therefore, the restriction of using an observer and its associated estimator as an estimating device is the same as restricting ones attention to only unbiased state estimators.

But if the a priori assumption on $\underline{x}_o$ is different from the true mean of $\underline{x}(0)$, then it is not unbiased minimum mean square error approach. In

fact in this situation, nearly all the approaches mentioned above may not be justified. But we shall see in discussions that under some mild conditions our approach is still valid even with incorrect on a priori distribution on the initial state $\underline{x}(0)$.

(B) Dimension of Observers

From the point of memory storage, we would like to find the minimal order optimum observer; but from the point of view of computation, one may not want to find the minimal order optimum observer.[†] One may want to look for those observers $\mathcal{G}_T$, $T \in \mathcal{J}_{V^*}$ where the number of nonzero entries of $\underline{T}(k+1)\ \underline{A}(k)\ \underline{P}(k)$, $\underline{T}(k+1)\ \underline{A}(k)\ \underline{V}(k)$, and $\underline{P}(k)$ is kept to a minimum. No systematic way of finding such observers is available; in general this will depend on the specific problem under consideration.

(C) Detectability and Observability

Detectability is a weaker condition than observability (see theorem 3.5.4). Essentially, detectability implies that in noise free situations, one can deduce the current state (but not the initial state) of the system if given infinitely long observation period, and so it is not the same as "asymptotic observability" (if such a concept can be defined). In all sequential estimation problems, one is interested to estimate the current state rather than the initial state of the system, so one would expect that detectability would be the intrinsic property which will assume nice behavior of the minimum error covariance when noises are present. This physical intuition was verified in section 3.6. We showed that detectability of linear system gives the necessary and sufficient condition for

---

[†]This viewpoint is due to F.C. Schweppe

uniformly bounded $\underline{\Sigma}^*(k)$ and the existence of its steady state behavior. Observability implies that in the noise free situation, we can deduce the initial state of the system if given a long enough observation period. Of course, knowing the initial state will enable us to deduce the current state; but as long as sequential estimation is the goal, the knowledge of initial state will be nice but not absolutely necessary. Except in the smoothing estimation, where we are interested in finding not only the current estimate, but the whole trajectory estimate; thus in this situation, detectability may not be enough to assure the "nice behavior" of $\underline{\Sigma}^*(k)$; we need observability of the system.

In the development, we assume an a priori distribution on the initial state $\underline{x}(0)$. This assumption can only be justified if, as time advances and information accumulates, the resulting performance will be independent of the a priori distribution of $\underline{x}(0)$. Assume that the true mean of $\underline{x}(0)$ is $\underline{x}_o'$ but we guess its mean to be $\underline{x}_o$. Since the mean of the state of S_2 satisfies (a.s.) the deterministic equation described by S_1 (see section 2.3), then detectability implies that even with a wrong assumption on the mean of $\underline{x}(0)$, the optimum observer will give an asymptotically unbiased estimate Thus as $k \to \infty$, $\underline{\Sigma}^*(k)$ truly represents the error covariance. From theorem 3.6.3, we see that in the steady state period, the error covariance is independent of the covariance of $\underline{x}(0)$. Therefore if S_2 is detectable, the performance will "merge" when information accumulates even if we started off with different assumptions on the statistical law $\underline{x}(0)$. Thus detectability justifies the assumption on knowing the mean and covariance of $\underline{x}(0)$.

(D) Sequentially-Correlated Observation Noise

The derived results are also applicable to the case when the observation noise satisfies: [see equation (3.3.1) for S]

$$\underline{\eta}(k+1) = \bar{A}(k)\underline{\eta}(k) + \underline{\gamma}(k) \quad ; \quad \underline{\eta}(k) \in R^m \qquad (3.7.13)$$

where $\{\underline{\gamma}(k)\}_{k=0}^{\infty}$, $\underline{\eta}(0)$, $\underline{x}(0)$, and $\{\underline{\xi}(k)\}_{k=0}^{\infty}$ are independent with statistical law (3.3.2), (3.3.3) and

$$\underline{\eta}(0) \sim G(\underline{\eta}_o, \Sigma_o^{\eta}) \quad ; \quad \underline{\gamma}(k) \sim G(\underline{0}, \bar{Q}(k)) \quad . \qquad (3.7.14)$$

We can define

$$\underline{x}^a(k) \triangleq \begin{bmatrix} \underline{x}(k) \\ \cdots \\ \underline{\eta}(k) \end{bmatrix} \; ; \; \underline{A}^a(k) = \begin{bmatrix} \underline{A}(k) & \vdots & \underline{0} \\ \cdots & \vdots & \cdots \\ \underline{0} & \vdots & \bar{A}(k) \end{bmatrix} \; ;$$

$$\underline{\xi}^a(k) = \begin{bmatrix} \underline{\xi}(k) \\ \cdots \\ \underline{\gamma}(k) \end{bmatrix} \; ; \; \underline{B}^a(k) = \begin{bmatrix} \underline{B}(k) \\ \cdots \\ \underline{0} \end{bmatrix} \quad . \qquad (3.7.15)$$

Then we have the augmented system

$$S_2^a: \quad \begin{aligned} \underline{x}^a(k+1) &= \underline{A}^a(k) + \underline{B}^a(k)\underline{u}(k) + \underline{\xi}^a(k) \\ \underline{y}(k) &= [\underline{C}(k) \vdots \underline{I}_m]\underline{x}^a(k) \end{aligned} \quad . \qquad (3.7.16)$$

We can apply the derived results to the above system S_2^a. Note that $\underline{x}^a(k+1) \in R^{n+m}$, but since the noise free observation is of dimension m, the minimum order optimum observer is of order n. This problem has also been considered by Henrikson [46], Bryson and Ho [43] using a different approach. We can easily verify that the results obtained by applying the

derived results to this special class of problem ; they are the same as those obtained by Henrikson. This special application will be considered in a future investigation.

3.8 Perspective

Observers for a linear system were introduced by Luenberger [35], [36]. He only considered continuous,linear,time invariant systems. Observers for discrete,linear,time invariant systems were discussed by Aoki and Huddle [37] in relation to a constrained estimator problem. Observers for discrete linear time varying system were first introduced and studied by Tse and Athans [38].

Optimum linear filtering for discrete linear time varying systems was investigated by Kalman [39], [40] using the projection theorem approach. Deadbeat estimator for discrete time invariant system were derived by Kalman [41]. The unbiased approach to optimum linear filtering problems was used by Athans and Tse [10], Tse and Athans [38]; the unbiased approach to nonlinear filter was used by Athans, Wishner, Bertolini [42].

Detectability was first introduced by Wonham [32] as the dual concept of stabilizability. Detectability as defined by definition 3.5.1 seems to be more appropriate and more general than that of Wonham's (Wonham considered only the time invariant case).

The asymptotic behavior of minimum error covariance for discrete linear systems were not investigated in full detail in the current literature. Deyst and Price [28], Sorenson [29], and Aoki [30] considered to some extent the asymptotic properties of the minimum error covariance. They confine themselves to consider the special case when the observation noise is regular ($\underline{Q}(k) > \underline{0}$). Little or no attention is paid to the case when the

observation noise is degenerate ($\underline{Q}(k) \geq \underline{0}$) or singular ($\underline{Q}(k) = \underline{0}$). The treatment in section 3.6 is original, and considers all different cases in a unifying manner.

CHAPTER IV

OBSERVER THEORY FOR CONTINUOUS TIME LINEAR SYSTEMS

4.1 Introduction

The problem of state estimation for discrete linear systems was considered in detail in Chapter III. In this chapter, we shall consider the state estimation problem for continuous time linear dynamical systems. Aside from the fact that state estimation is of prime importance in the design of optimal control systems, the problem in itself is of great importance in the design of modern communication systems.

The structure of this chapter is as follows. In section 4.2, we consider time-varying deterministic linear systems; the notion of a deterministic observer and estimator for a continuous linear system is defined and we prove that classes of observers and estimators can be constructed if the dynamics of the system are known. Equivalent classes of observers and the classes of minimal order observers are defined and some preliminary results on parameterizing equivalent classes of observers are obtained. In section 4.3, we extend the deterministic notions to stochastic systems where we show that some classes of observers yield unbiased estimates. By some physical considerations, we restrict the classes of observer-estimators that shall be considered. We then determine the class of minimal order observers that yield minimum variance estimates by formulating the problem as finding the minimal function of a certain restricted solution set and then using theorem 2.6.3. We then show that the class of minimal order optimum observer-estimators yields the conditional mean estimates of the stage. This reveals the true nature of the derived minimal order optimum

observer-estimator. In section 4.4, the notion of detectability of continuous linear time system is defined, and the asymptotic behavior of the optimum estimator is studied in terms of detectability and observability of the system. In section 4.5, we have some general discussions on the approaches, results and further applications. In section 4.6, detailed literature connected with the development in this chapter is listed.

Conceptually, there is little difference between discrete and continuous time linear systems; therefore we would expect the results obtained in this chapter will be quite similar to those of Chapter III. One marked difference between the discrete and continuous time cases is that for the discrete time case, the observation statistic is sequential, and so each bit of observation conveys finite amount of information in an accumulative manner; whereas in the continuous time case, we have only a priori information before any observation is made, and when an observation is made at the initial time we have a sudden increase of information within a very small interval of time due to some noise-free observation component. We would expect this "jump" in information to be reflected in the initial condition of the optimum observer-estimator.

4.2 Classes of Observers for Continuous Linear Systems

In this section, we shall consider a linear time-varying continuous system S_1^c described by

(state eq.) $$\dot{\underline{x}}(t) = \underline{A}(t)\underline{x}(t) + \underline{B}(t)\,\underline{u}(t)$$

(output eq.) $$\underline{y}(t) = \underline{C}(t)\underline{x}(t)$$

where $\underline{x}(t) \in R^n$. We shall assume that $\underline{C}(t)$ is a differentiable time-varying nxm matrix of rank m, for all $t \in [t_o,\infty]$ $(n \geq m)$. For a fixed $t \in [t_o,\infty]$ the set of complementary matrices of order s for $\underline{C}(t)$ is denoted by $\Omega(\underline{C}(t);m,s,n) = \{\underline{T}(t) \in M_{sn}: N(\underline{T}(t)) \cap N(\underline{C}(t)) = \underline{0}_n \in R^n\}$. We note that $\underline{T}(t) \in \Omega(\underline{C}(t);m,s,n)$ if and only if there exist matrices $\underline{P}(t)$, $\underline{V}(t)$ (of appropriate dimensions) such that

$$\underline{P}(t)\underline{T}(t) + \underline{V}(t)\underline{C}(t) = \underline{I}_n \quad . \tag{4.2.1}$$

Definition 4.2.1: A linear time varying system of dimension $s \geq n - m$

$$\Theta_s: \quad \dot{\underline{z}}(t) = \underline{F}(t)\underline{z}(t) + \underline{D}(t)\underline{y}(t) + \underline{G}(t)\underline{u}(t) \quad ; \quad \underline{z}(t_o^+) = \underline{z}_o \tag{4.2.2}$$

is an s-order observer for the system $\mathbb{S}$, if for some choices of $\underline{z}_o$, the solution, $\underline{z}(t)$ of (4.2.2) equals

$$\underline{z}(t) = \underline{T}(t)\underline{x}(t) \quad ; \quad t > t_o \tag{4.2.3}$$

for some $\underline{T}(t) \in \Omega(\underline{C}(t);m,s,n)$, $t > t_o$. We shall also say that the observer is described by $\underline{T}(t)$, $t > t_o$; and refer to such an observer by the symbol Θ_T^c.

Let $\underline{T}(t)$ be an sxn matrix which satisfies the differential equation $(t > t_o)$

$$\dot{\underline{T}}(t) = \underline{F}(t)\underline{T}(t) - \underline{T}(t)(\underline{A}(t) - \underline{L}(t)\underline{C}(t) + \tilde{\underline{D}}(t)\underline{C}(t) \quad ; \quad \underline{T}(t_o^+) = \underline{T}_o \tag{4.2.4}$$

where $\underline{L}(t)$, $\underline{F}(t)$, $\tilde{\underline{D}}(t)$, $\underline{T}_o$ are some prescribed matrices of appropriate dimensions. If we construct a time varying system of dimension $s \geq n - m$:

$$\Theta_s': \dot{\underline{z}}(t) = \underline{F}(t)z(t) + (\tilde{\underline{D}}(t) + \underline{T}(t)\underline{L}(t))\underline{y}(t) + \underline{T}(t)\underline{B}(t)\underline{u}(t) \tag{4.2.5}$$

then using (4.2.4), we have

$$\frac{d}{dt}\left(\underline{T}(t)\underline{x}(t) - \underline{z}(t)\right) = \underline{F}(t)\left(\underline{T}(t)\underline{x}(t) - \underline{z}(t)\right) \tag{4.2.6}$$

and if we choose $\underline{T}_o\ \underline{x}(t_o) = \underline{z}(t_o^+)$, then $\underline{T}(t)\underline{x}(t) = \underline{z}(t)$, $t > t_o$.

Therefore $\mathcal{O}_s'$ will be an s-order observer for $\mathcal{S}_1^c$ if by some appropriate choices of $\underline{L}(t)$, $\underline{F}(t)$, $\underline{D}(t)$, $\underline{T}_o$ the solution, $\underline{T}(t)$ of (4.2.4) will be in the set of complementary matrices of order s for $\underline{C}(t)$, $t > t_o$.

By assumption, $\underline{C}(t)$ is differentiable for all $t \in [t_o,\infty]$; thus there exists a function $\tilde{\underline{T}}(t) \in \mathcal{L}(\underline{C}(t);m,s,n)$, $t > t_o$, such that $\tilde{\underline{T}}(t)$ is differentiable.

<u>Theorem 4.2.2</u>: Let $\tilde{\underline{T}}(t) \in \mathcal{L}(\underline{C}(t);m,s,n)$, $t > t_o$, and $\tilde{\underline{T}}(t)$ is differentiable in the interval (t_o,∞). Then, there exists a class of s-order observers which are all described by $\tilde{\underline{T}}(t)$, $t > t_o$ for the system $\mathcal{S}_1^c$.

<u>Proof</u>: Let $\tilde{\underline{P}}(t)$, $\tilde{\underline{V}}(t)$, $t > t_o$, be matrices of appropriate dimension such that

$$\tilde{\underline{P}}(t)\tilde{\underline{T}}(t) + \tilde{\underline{V}}(t)\underline{C}(t) = \underline{I}_n \quad ; \quad t > t_o \quad . \tag{4.2.7}$$

Choose for $t > t_o$

$$\underline{D}(t) = \tilde{\underline{T}}(t)\tilde{\underline{A}}(t)\tilde{\underline{V}}(t) + \dot{\tilde{\underline{T}}}(t)\tilde{\underline{V}}(t) \tag{4.2.8}$$

$$\underline{F}(t) = \tilde{\underline{T}}(t)\tilde{\underline{A}}(t)\tilde{\underline{P}}(t) + \dot{\tilde{\underline{T}}}(t)\tilde{\underline{P}}(t) \tag{4.2.9}$$

$$\underline{T}_o = \tilde{\underline{T}}(t_o^+) \tag{4.2.10}$$

where

$$\tilde{\underline{A}}(t) \triangleq \underline{A}(t) - \underline{L}(t)\underline{C}(t) \tag{4.2.11}$$

and $\underline{L}(t)$ is an arbitrary n×m matrix. With these choices of $\underline{D}(t)$, $\underline{F}(t)$, and $\underline{T}_c$, we have the solution of (4.2.4)

$$\underline{T}(t) = \underline{\Phi}_F(t,t_o^+)\tilde{\underline{T}}(t_o^+)\underline{\Phi}_{\tilde{A}}(t_o^+,t)+\int_{t_o}^{t}\underline{\Phi}_F(t,\tau)[\tilde{\underline{T}}(\tau)\tilde{\underline{A}}(\tau)+\frac{d\underline{T}(\tau)}{d\tau}\tilde{\underline{V}}(\tau)\underline{C}(\tau)]\underline{\Phi}_{\tilde{A}}(\tau,t)d\tau \tag{4.2.12}$$

where $\underline{\Phi}_F(t,t_o^+)$ and $\underline{\Phi}_{\tilde{A}}(t,t_o^+)$ are fundamental matrices associated with $\underline{F}(t)$ and $\tilde{\underline{A}}(t)$ respectively. Using (4.2.7) and (4.2.9) the integrand of (4.2.12) becomes

$$\int_{t_o}^{t}\underline{\Phi}_F(t,\tau)[\tilde{\underline{T}}(\tau)\tilde{\underline{A}}(\tau)+\frac{d\tilde{\underline{T}}(\tau)}{d\tau}\tilde{\underline{V}}(\tau)\underline{C}(\tau)]\underline{\Phi}_{\tilde{A}}(\tau,t)d\tau = \int_{t_o}^{t}[\underline{\Phi}_F(t,\tau)\tilde{\underline{T}}(\tau)\frac{d\underline{\Phi}_{\tilde{A}}(\tau,t)}{d\tau}+$$

$$\underline{\Phi}_F(t,\tau)\frac{d\tilde{\underline{T}}(\tau)}{d\tau}\underline{\Phi}_{\tilde{A}}(\tau,t)+\frac{d\underline{\Phi}_F(t,\tau)}{d\tau}\tilde{\underline{T}}(\tau)\underline{\Phi}_{\tilde{A}}(\tau,t)]d\tau = \tilde{\underline{T}}(t)-\underline{\Phi}_F(t,t_o^+)\tilde{\underline{T}}(t_o^-)\underline{\Phi}_{\tilde{A}}(t_o^+,t) \tag{4.2.13}$$

Combining (4.2.12) and (4.2.13) we have

$$\underline{T}(t) = \tilde{\underline{T}}(t) \in \Omega(\underline{C}(t);m,s,n) \quad , \quad t > t_o \quad . \tag{4.2.14}$$

So an s-order observer can be constructed by (4.2.5). We note that by choosing different $\underline{L}(t) \in M_{nm}$, $t > t_o$, we obtain a class of observers described by the same $\tilde{\underline{T}}(t)$, $t > t_o$. For a fixed $\tilde{\underline{T}}(t) \in \Omega(\underline{C}(t);m,s,n)$, $t > t_o$, and a fixed $\underline{L}(t) \in M_{nm}$, $t > t_o$, we shall use the symbol $\mathcal{O}_{\tilde{T}}^{1c}(L)$ to represent the observer which is specified by $\tilde{\underline{T}}(t)$ and $\underline{L}(t)$, and $\mathcal{O}_{\tilde{T}}^{1c} = \{\mathcal{O}_{\tilde{T}}^{1c}(L)/\underline{L}(t) \in M_{nm}\}$ the class of observers which is specified by $\tilde{\underline{T}}(t)$.

For each $\mathcal{O}_{\tilde{T}}^{1c}(L) \in \mathcal{O}_{\tilde{T}}^{1c}$, we shall associate with it an estimator $\mathcal{E}_{\tilde{T}}^{1c}(L)$ described by (Figure 4.1)

$$\mathcal{E}_{\tilde{T}}^{1c}(L): \begin{array}{l} \dot{\underline{z}}(t) = (\tilde{\underline{T}}(t)(\underline{A}(t)-\underline{L}(t)\underline{C}(t))\tilde{\underline{P}}(t)+\dot{\tilde{\underline{T}}}(t)\tilde{\underline{P}}(t))\underline{z}(t)+(\tilde{\underline{T}}(t)(\underline{A}(t)-\underline{L}(t)\underline{C}(t)\tilde{\underline{V}}(t) \\ \qquad +\dot{\tilde{\underline{T}}}(t)\tilde{\underline{V}}(t)+\tilde{\underline{T}}(t)\underline{L}(t))\underline{y}(t)+\tilde{\underline{T}}(t)\underline{B}(t)\underline{u}(t) \\ \underline{w}(t) = \tilde{\underline{P}}(t)\underline{z}(t)+\tilde{\underline{V}}(t)\underline{y}(t) \end{array} \tag{4.2.15}$$

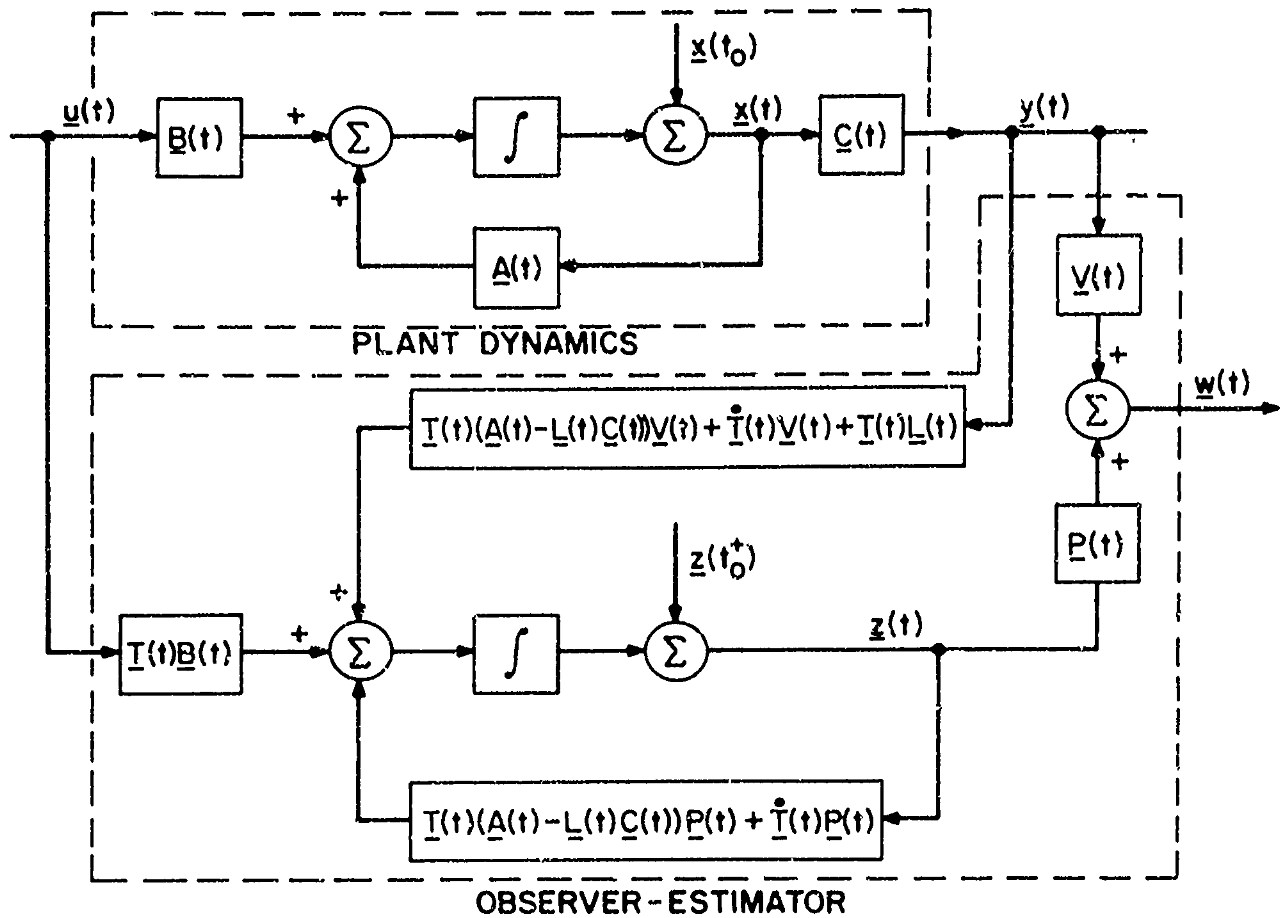


Fig. 4.1 THE STRUCTURE OF AN OBSERVER-ESTIMATOR FOR A DETERMINISTIC CONTINUOUS TIME LINEAR SYSTEM. THE VECTOR $\underline{w}(t)$ IS THE ESTIMATE OF THE STATE VECTOR $\underline{x}(t)$.

If we know $\underline{x}(t_o^+)$, then by setting

$$\underline{z}(t_o^+) = \tilde{\underline{T}}(t_o^+)\underline{x}(t_o^+) \tag{4.2.16}$$

we have from (4.2.6) and (4.2.16) that

$$\underline{w}(t) = \tilde{\underline{P}}(t)\tilde{\underline{T}}(t)\underline{x}(t) + \tilde{\underline{V}}(t)\underline{C}(t)\underline{x}(t) = \underline{x}(t) \quad . \tag{4.2.17}$$

But usually $\underline{x}(t_o^+)$ is unknown, and so if we want to use $\mathcal{E}_{\tilde{T}}^{lc}(L)$ as an estimating device, we would restrict the initial condition of $\underline{z}(t_o^+)$ to be in the range space of $\tilde{\underline{T}}(t_o^+)$.

Let $\underline{V}(t) \in \mathcal{M}_{nm}$, $t > t_o$, be a fixed differentiable matrix. Associated with it is a set of matrix functions $\mathcal{J}_V^c = \{\underline{T}(t) \in \mathcal{M}_{sn}, t > t_o / \underline{T}(t)$ is differentiable on (t_o,∞) and $\underline{P}(t)\underline{T}(t) + \underline{V}(t)\underline{C}(t) = \underline{I}_n$ for some $\underline{P}(t) \in \mathcal{M}_{ns}$, $t \in (t_o,\infty); s \geq n - m\}$. For a fixed $\underline{T}(t) \in \mathcal{J}_V^c$ we can associate with it a class of observers $\mathcal{O}_T^{lc}$ and a class of estimators $\mathcal{E}_T^{lc} = \{\mathcal{E}_T^{lc}(L)/\underline{L}(t) \in \mathcal{M}_{nm}\}$. Therefore, for a fixed $\underline{V}(t)$,* we can associate with it different classes of observers $\mathcal{O}_T^{lc}$. $\underline{T}(t) \in \mathcal{J}_V^c$, of different orders.

For a fixed $\underline{V}(t)$, $t > t_o$ suppose that $(\underline{I}_n - \underline{V}(t)\underline{C}(t))$ has rank $n - p$, $(p \leq m)$: then the class $\mathcal{F}_V^{pc}(L) = \{\mathcal{O}_T^{lc}(L)/\underline{T}(t) \in \mathcal{J}_V^c$ and $\underline{T}(t) \in \mathcal{M}_{n(n-p)}$, has full rank, $t \in (t_o,\infty)\}$ is called the class of minimal order observers associated with $\underline{V}(t)$ and parameterized by $\underline{L}(t)$. We can define the notion of equivalent representation as in the discrete case (Definition 3.2.5).

---

*For the rest of this chapter, $\underline{V}(t)$ is always assumed to be differentiable on $t \in (t_o,\infty]$.

Lemma 4.2.3: For a fixed $\underline{V}(t)$, $t \geq t_o$ such that $\underline{I}_n - \underline{V}(t)\underline{C}(t)$ has rank $n - p$, $t \geq t_o$, let $\mathcal{O}_T^{1c}(L)$, $\underline{T}(t) \in \mathcal{T}_V^c$ be a given observer of order $s \geq n - p$. Then there exists a $n - p$ order observer $\mathcal{O}_{\tilde{T}}^{1c}(L)$, $\underline{\tilde{T}}(t) \in \mathcal{T}_V^c$ such that $\mathcal{O}_{\tilde{T}}^{1c}(L)$, $\mathcal{O}_T^{1c}(L)$ are equivalent.

<u>Proof</u>: Let $\mathcal{O}_T^{1c}(L)$ be a given s-order observer; its associated estimator $\mathcal{E}_T^{1c}(L)$ is described by

$$
\mathcal{E}_T^{1c}(L): \quad
\begin{aligned}
\underline{\dot{z}}(t) &= (\underline{T}(t)(\underline{A}(t)-\underline{L}(t)\underline{C}(t)\underline{P}(t)+\underline{\dot{T}}(t)\underline{P}(t))z(t)+(\underline{T}(t)(\underline{A}(t)-\underline{L}(t)\underline{C}(t) \\
&\qquad \underline{V}(t)+\underline{\dot{T}}(t)\underline{V}(t)\underline{y}(t)+\underline{T}(t)\underline{B}(t)\underline{u}(t) \quad ; \quad \underline{z}(t_o^+)\in S = \{\underline{T}(t_o^+)\underline{x} \mid \underline{x}\in R^n\} \\
\underline{w}(t) &= \underline{P}(t)\underline{z}(t)+\underline{V}(t)\underline{y}(t)
\end{aligned}
\tag{4.2.18}
$$

with $\underline{P}(t)$, $\underline{T}(t)$ satisfying (4.2.1) and $\underline{z}(t) \in R^s$, $s \geq n - p$. Since $\underline{I}_n - \underline{V}(t)\underline{C}(t)$ has rank $n - p$, we may assume without loss of generality that $\underline{P}(t)$ is of rank $n - p$. $\underline{P}(t)$ is a time varying linear transformation from $R^s \to R^n$. We can break the transformation into two steps: map R^s to R^{n-p} by a time invariant transformation $\underline{K}$, then from R^{n-p} to R^n by an appropriate time varying transformation $\underline{\tilde{P}}(t)$ i.e.,

$$
\underline{P}(t) = \underline{\tilde{P}}(t)\underline{K} \qquad ; \qquad \underline{\tilde{P}}(t) \in M_{n(n-p)} \qquad \underline{K} \in M_{(n-p)s} \,. \tag{4.2.19}
$$

Let us construct an $n - p$ order observer $\mathcal{O}_{\tilde{T}}^{1c}(L)$ with $\underline{\tilde{T}}(t) = \underline{K}\ \underline{T}(t)$ and the restricted observer's state initial condition $\underline{\tilde{z}}(t_o^+) \in \tilde{S} = \{\underline{K}\ \underline{T}(t_o^+)\underline{x} \mid \underline{x} \in R^n\}$. First we see from (4.2.19) that

$$
\underline{\tilde{P}}(t)\underline{\tilde{T}}(t)+\underline{V}(t)\underline{C}(t) = \underline{\tilde{P}}(t)\underline{K}\ \underline{T}(t)+\underline{V}(t)\underline{C}(t) = \underline{P}(t)\underline{T}(t)+\underline{V}(t)\underline{C}(t) = \underline{I}_n \tag{4.2.20}
$$

thus we conclude that $\underline{\tilde{T}}(t) \in \mathcal{T}_V^c$. Let $\mathcal{E}_{\tilde{T}}^{1c}(L)$ be the estimator associated with $\mathcal{O}_{\tilde{T}}^{1c}(L)$. To prove the lemma, we need to verify that $\underline{\tilde{w}}(t) = \underline{w}(t)$ for

all possible $\underline{u}(t)$ and $\underline{y}(t)$ where $\tilde{\underline{w}}(t)$ is the output of $\mathcal{E}^{1c}_{\tilde{T}}(L)$. Let $\underline{z}(t_o^+) = \underline{T}(t_o^+)\underline{\alpha}$ for some $\underline{\alpha} \in R^n$ and pick $\tilde{\underline{z}}(t_o^+) = \underline{K}\ \underline{T}(t_o^+)\underline{\alpha}$; then one can easily show that by construction

$$\tilde{\underline{z}}(t) = \underline{K}\ \underline{z}(t) \quad ; \quad t > t_o \tag{4.2.21}$$

for all $\underline{y}(t)$ and $\underline{u}(t)$. Then we have

$$\tilde{\underline{w}}(t) = \tilde{\underline{P}}(t)\tilde{\underline{z}}(t)+\underline{V}(t)\underline{y}(t) = \tilde{\underline{P}}(t)\underline{K}\ \underline{z}(t)+\underline{V}(t)\underline{y}(t) = \underline{P}(t)\underline{z}(t)+\underline{V}(t)\underline{y}(t) = \underline{w}(t) \ . \tag{4.2.22}$$

Conversely, if $\tilde{\underline{z}}(t_o^+) = \underline{K}\ \underline{T}(t_o^+)\underline{\alpha}$, pick $\underline{z}(t_o^+) = \underline{T}(t_o^+)\underline{\alpha}$; then we have (4.2.21) and (4.2.22) in the same manner, and the lemma is proved.

<u>Theorem 4.2.4</u>: Let $\underline{V}(t) \in M_{nm}$, $t > t_o$, such that the rank of $\underline{V}(t)$ and $\underline{I}_n - \underline{V}(t)\underline{C}(t)$ are p and n - p, respectively. For a fixed $\underline{L}(t) \in M_{nm}$ the class of observers $\mathfrak{S}^{1c}_T(L)$, $\underline{T}(t) \in T^c_V$ are all equivalent.

<u>Proof</u>: Let $\underline{P}(t) \in M_{n(n-p)}$, $\underline{T}(t) \in M_{(n-p)n}$ such that

$$\underline{P}(t)\underline{T}(t) + \underline{V}(t)\underline{C}(t) = [\underline{P}(t) \ \vdots \ \underline{V}(t)]\begin{bmatrix}\underline{T}(t) \\ \cdots \\ \underline{C}(t)\end{bmatrix} = \underline{I}_n \ . \tag{4.2.23}$$

Denote the column vectors of $\underline{V}(t)$ by $\underline{v}_i(t)$, i = 1, ..., m:

$$\underline{V}(t) = \begin{bmatrix} \updownarrow & \updownarrow & & \updownarrow \\ \underline{v}_1(t) & \underline{v}_2(t) & \cdots & \underline{v}_m(t) \\ \updownarrow & \updownarrow & & \updownarrow \end{bmatrix} \ . \tag{4.2.24}$$

Since $\underline{V}(t)$ has rank p, $\{\underline{v}_{\sigma_t(i)}(t)\}_{i=1}^p$ form an independent set, where $\sigma_t(\cdot)$ is a permutation of 1, ..., m and $\underline{v}_{\sigma_t(j)}$, j > p are dependent on $\{\underline{v}_{\sigma_t(i)}(t)\}_{i=1}^p$. Rearranging, if necessary, we may assume for a fixed t:

$$\underline{V}(t) = [\underline{V}_1(t) \ \vdots \ \underline{V}_2(t)] \tag{4.2.25}$$

with $\underline{V}_1(t) \in M_{np}$ and of rank p, while

$$\underline{V}_2(t) = \underline{V}_1(t)\underline{M}(t) \quad ; \quad \underline{M}(t) \in M_{p(m-p)} \quad . \qquad (4.2.26)$$

The matrix $\underline{C}(t)$ is also rearranged accordingly (if necessary); we may assume

$$\underline{C}(t) = \begin{bmatrix} \underline{C}_1(t) \\ \cdots \\ \underline{C}_2(t) \end{bmatrix} \quad ; \quad \underline{C}_1(t) \in M_{pn} \quad , \quad \underline{C}_2(t) \in M_{(n-p)n} \qquad (4.2.27)$$

Since $\underline{C}(t)$ is of full rank, (4.2.23) implies that

$$\underline{C}_2(t) = \underline{N}(t)\underline{T}(t) \quad ; \quad \underline{N}(t) \in M_{(m-p)(n-p)} \quad . \qquad (4.2.28)$$

Using (4.2.23) to (4.2.28), we have for a fixed t:

$$[\underline{P}(t) \vdots \underline{V}_1(t)] \begin{bmatrix} \underline{T}(t) \\ \cdots \\ \underline{C}_1(t) + \underline{M}(t)\underline{N}(t)\underline{T}(t) \end{bmatrix} = \underline{I}_n \quad . \qquad (4.2.29)$$

Since $[\underline{P}(t) \vdots \underline{V}_1(t)] \in M_{nn}$, (4.2.29) implies that

$$\left.\begin{aligned} &\underline{T}(t)\underline{P}(t) = \underline{I}_{n-p} \quad ; \quad \underline{T}(t)\underline{V}_1(t) = \underline{0}_{(n-p)p} \quad ; \\ &\underline{C}_1(t)\underline{V}_1(t) + \underline{M}(t)\underline{N}(t)\underline{T}(t)\underline{V}_1(t) = \underline{C}_1(t)\underline{V}_1(t) = \underline{I}_p \quad . \end{aligned}\right\} \qquad (4.2.30)$$

From (4.2.25), (4.2.26) and (4.2.30), we have

$$\underline{T}(t)\underline{P}(t) = \underline{I}_{n-p} \quad ; \quad \underline{T}(t)\underline{V}(t) = \underline{0}_{(n-p)m} \quad . \qquad (4.2.31)$$

We note that under the assumption on $\underline{V}(t)$, (4.2.31) is true for all $t \geq t_o$.

From lemma 4.2.3, we see that to prove the theorem we need only to prove that all observers $\Theta_T^{lc}(L) \in \pi_V^{pc}(L)$ are equivalent.

Let $\Theta_{T_i}^{lc}(L) \in \pi_V^{pc}(L)$ be arbitrary, i = 1, 2. The associated estimators are described by

$$\mathcal{E}^{1c}_{T_i}(L): \quad \begin{aligned} \dot{\underline{z}}_i(t) &= (\underline{T}_i(t)(\underline{A}(t)-\underline{L}(t)\underline{C}(t))\underline{P}_i(t)+\dot{\underline{T}}_i(t)\underline{P}_i(t))\underline{z}_i(t)+(\underline{T}_i(t)\underline{L}(t)+\dot{\underline{T}}_i(t)\underline{V}(t) \\ &\quad +\underline{T}_i(t)(\underline{A}(t)-\underline{L}(t)\underline{C}(t))\underline{V}(t))\underline{y}(t)+\underline{T}_i(t)\underline{B}(t)\underline{u}(t) \\ \underline{w}_i(t) &= \underline{P}_i(t)\underline{z}_i(t)+\underline{V}(t)\underline{y}(t) \quad ; \quad \underline{z}_i(t_o^+)\varepsilon S_i = \{\underline{T}_i(t_o^+)\underline{\alpha} \mid \underline{\alpha}\varepsilon R^n\} \end{aligned} \tag{4.2.32}$$

Without loss of generality, we may assume that the $\underline{P}_i(t) \varepsilon M_{n(n-p)}$ are of rank n - p. Then there exists a nonsingular matrix $\underline{K}(t) \varepsilon M_{(n-p)(m-p)}$, such that

$$\underline{P}_1(t) = \underline{P}_2(t)\underline{K}(t) \quad ; \quad \underline{P}_2(t) = \underline{P}_1(t)\underline{K}^{-1}(t) \tag{4.2.33}$$

and so we also have

$$\underline{K}(t)\underline{T}_1(t) = \underline{T}_2(t) \quad ; \quad \underline{T}_1(t) = \underline{K}^{-1}(t)\underline{T}_2(t) \quad . \tag{4.2.34}$$

Let us define

$$\tilde{\underline{z}}(t) = \underline{K}(t)\underline{z}_1(t) \quad . \tag{4.2.35}$$

Using (4.2.32) to (4.2.35), we obtain the equation for $\tilde{\underline{z}}(t)$:

$$\begin{aligned} \dot{\tilde{\underline{z}}}(t) &= (\underline{T}_2(t)(\underline{A}(t)-\underline{L}(t)\underline{C}(t))\underline{P}(t)+\underline{K}(t)\dot{\underline{T}}_1(t)\underline{P}_2(t)+\dot{\underline{K}}(t)\underline{K}^{-1}(t))\tilde{\underline{z}}(t) \\ &\quad +(\underline{T}_2(t)\underline{L}(t)+\underline{K}(t)\dot{\underline{T}}_1(t)\underline{V}(t)+\underline{T}_2(t)(\underline{A}(t)-\underline{L}(t)\underline{C}(t))\underline{V}(t))\underline{y}(t) \\ &\quad +\underline{T}_2(t)\underline{B}(t)\underline{u}(t) \; ; \\ \tilde{\underline{z}}(t_o^+) &= \underline{K}(t_o^+)\underline{T}_1(t_o^+)\underline{\alpha} = \underline{T}_2(t_o^+)\underline{\alpha} \;\varepsilon\; S_2 \end{aligned} \tag{4.2.36}$$

since $\underline{P}_i(t)$, $\underline{T}_i(t)$ satisfy (4.2.23), i = 1, 2; thus, by (4.2.31), (4.2.33) and (4.2.34) we can easily show that

$$\underline{K}(t)\dot{\underline{T}}_1(t)\underline{P}_2(t) + \dot{\underline{K}}(t)\underline{K}^{-1}(t) = \dot{\underline{T}}_2(t)\underline{P}_2(t) \tag{4.2.37}$$

$$\underline{K}(t)\dot{\underline{T}}_1(t)\underline{V}(t) = \underline{K}(t)\dot{\underline{T}}_2(t)\underline{V}(t) \tag{4.2.38}$$

Substituting (4.2.37), (4.2.38) into (4.2.36) and comparing with (4.2.32), we see for any given $\underline{\tilde{z}}(t_o^+)$, we can pick an appropriate $\underline{z}_2(t_o^+) \varepsilon S_2(\underline{\tilde{z}}(t_o^+) = \underline{z}_2(t_o^+))$ such that

$$\underline{\tilde{z}}(t) = \underline{z}_2(t) \tag{4.2.39}$$

and so

$$\underline{w}(t) = \underline{P}_1(t)\underline{z}_1(t) + \underline{V}(t)\underline{y}(t) = \underline{P}_2(t)\underline{K}(t)\underline{z}_1(t) + \underline{V}(t)\underline{y}(t)$$

$$= \underline{P}_2(t)\underline{z}_2(t) + \underline{V}(t)\underline{y}(t) = \underline{w}_2(t) \quad . \tag{4.2.40}$$

Therefore $\mathcal{E}_{T_2}^{1c}(L)$ is an equivalent representation of $\mathcal{E}_{T_2}^{1c}(L)$; similarly, we can prove $\mathcal{E}_{T_1}^{1c}(L)$ is an equivalent representation of $\mathcal{E}_{T_2}^{1c}(L)$ and the theorem follows.

Note that the results are different from those in the discrete case. We see that only in some special cases equivalent classes of observers are parameterized by $\underline{V}(t) \varepsilon M_{nm}$ and $\underline{L}(t) \varepsilon M_{nm}$. Because of this difference, our approach to the problem of designing "nice-behaved" observers and associated estimatorsfor the continuous system $\mathcal{S}_1^c$ will be slightly different from that used in the discrete case.

4.3 Optimum Class of Observers for Linear Stochastic Systems

Consider a stochastic system $\mathcal{S}_2^c$ described by: (Figure 4.2)

$$\mathcal{S}_2^c: \begin{aligned} \dot{\underline{x}}(t) &= \underline{A}(t)\underline{x}(t) + \underline{B}(t)\underline{u}(t) + \underline{\xi}(t) \\ \underline{y}(t) &= \begin{bmatrix} \underline{C}_1(t)\underline{x}(t) + \underline{\eta}(t) \\ \cdots \\ \underline{C}_2(t)\underline{x}(t) \end{bmatrix} \end{aligned} \tag{4.3.1}$$

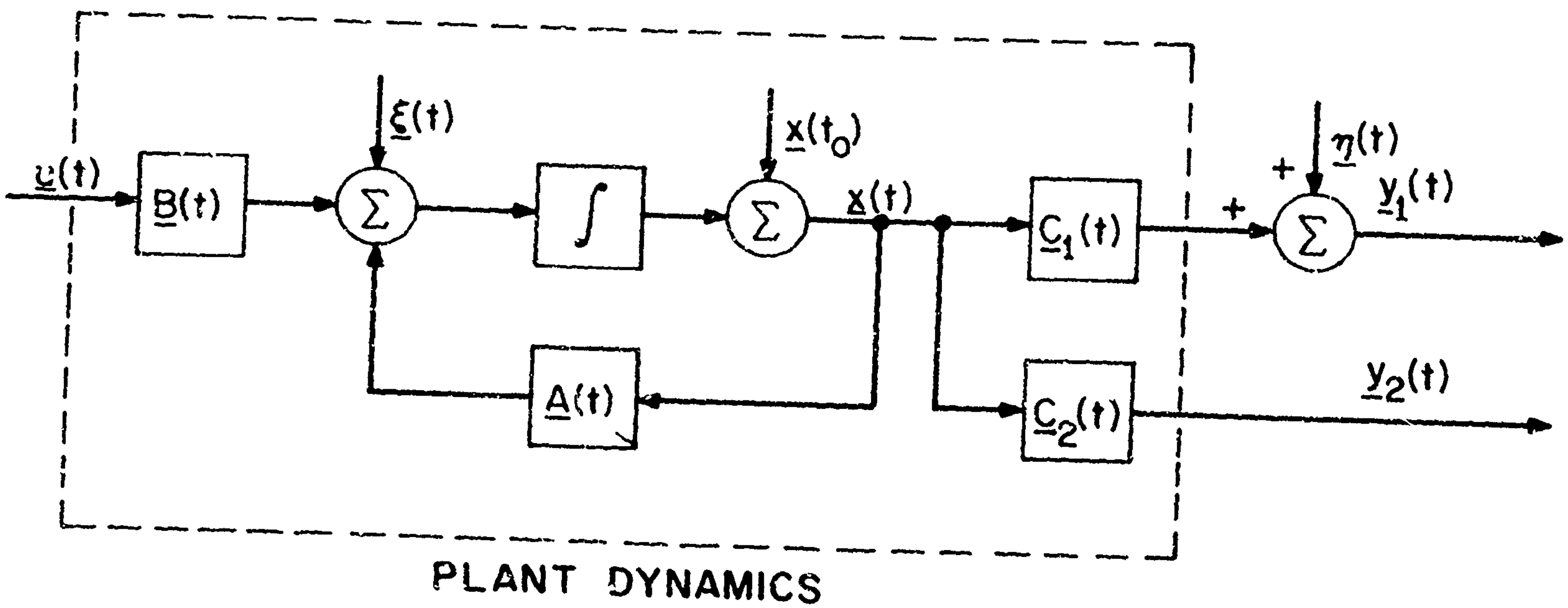


Fig. 4.2 THE STRUCTURE OF A CONTINUOUS TIME STOCHASTIC LINEAR SYSTEM S_2^c WITH SOME OBSERVATION CHANNELS NOISE-FREE

where $\underline{x}(t) \in R^n$, $\underline{u}(t) \in R^r$, $\underline{\xi}(t) \in R^n$, $\underline{\eta}(t) \in R^{m_1}$, $\underline{y}(t) \in R^m$, $(m \geq m_1)$. We assume that $\underline{x}(t_o)$, $\{\underline{\xi}(t),\ t \geq t_o\}$, $\{\underline{\eta}(t),\ t \geq t_o\}$ are independent statistics. $\underline{x}(t_o)$ is a Gaussian random vector with mean $\underline{x}_c$ and covariance $\underline{\Sigma}_o$; $\underline{\xi}(t)$, $\underline{\eta}(t)$, $t \geq t_o$ are white Gaussian noises with properties

$$E\left\{\int_{t_1}^{t_2} \underline{\xi}(t)dt\right\} = \underline{0} \quad ; \quad E\left\{\left(\int_{t_1}^{t_2} \underline{\xi}(t)dt\right)\left(\int_{t_1}^{t_2} \underline{\xi}(t)dt\right)'\right\} = \int_{t_1}^{t_2} \underline{R}(t)dt \quad ; \quad t_2 > t_1$$

$$E\left\{\int_{t_1}^{t_2} \underline{\eta}(t)dt\right\} = \underline{0} \quad ; \quad E\left\{\left(\int_{t_1}^{t_2} \underline{\eta}(t)dt\right)\left(\int_{t_1}^{t_2} \underline{\eta}(t)dt\right)'\right\} = \int_{t_1}^{t_2} \underline{Q}(t)dt \quad ; \quad t_2 > t_1 \tag{4.3.2}$$

where $\underline{R}(t) \geq \underline{0}$, $\underline{R}(t) \in M_{nn}$ and $\underline{Q}(t) > \underline{0}$, $\underline{Q}(t) \in M_{m_1 m_1}$. The control $\underline{u}(t)$ is known function of time.

Let us denote the noisy observation by $\underline{y}_1(t)$ and the noise-free observation by $\underline{y}_2(t)$:

$$\underline{y}_1(t) = \underline{C}_1(t)\underline{x}(t) + \underline{\eta}(t) \quad ; \quad \underline{y}_2(t) = \underline{C}_2(t)\underline{x}(t) \quad . \tag{4.3.3}$$

Our objective is to find a "filter" whose output will be an unbiased minimum mean square estimates of $\underline{x}(t)$. Since $\underline{x}(t)$ is a Gaussian random process (see Chapter 2, section 2.3), we may restrict ourselves to consider only linear filters [47]. Thus we may assume that the estimate of $\underline{x}(t)$ is given by

$$\mathcal{E}:\ \underline{w}(t) = \int_{t_o}^{t} \underline{H}(t,\tau)\underline{y}(\tau)d\tau + \underline{V}(t)\underline{y}(t) \tag{4.3.4}$$

where $\underline{H}(\cdot,\cdot)$ is an nxm matrix whose elements are differentiable in both arguments. If we demand the system $\mathcal{E}$ to yield unbiased estimates of $\underline{x}(t)$, then $\mathcal{E}$ can be realized by an s-order observer $\mathcal{O}_T^{2c}(L)^\dagger$ and its associated

†The superscript 2 is to indicate that system $\mathcal{S}_2^c$ is considered.

estimator $\mathcal{E}_T^{2c}(L)$ (see section 4.5). For this reason, we may view observer-estimators as estimating devices.

Let us restrict ourselves only to some special classes of observers. First we note that in the discrete analog, the optimum observers are compatible with respect to the noise-free observation. This is one intrinsic property of the optimum observers, and this property should be preserved when we pass from the discrete to the continuous. Thus, we shall only consider observers which are compatible† with respect to the noise-free observation $\underline{y}_2(t)$. Since $\underline{y}_1(t)$ contains white noise in the measurement, therefore, in order to obtain reasonable estimate, we shall not "pass" $\underline{y}_1(t)$ without filtering. These physical consideration allow us to consider only those observer $\mathcal{E}_T^{2c}(L)$ which are compatible and parameterized by $\underline{L}(t) \varepsilon M_{nm}$ arbitrary and $\underline{V}(t)$ of the form

$$\underline{V}(t) = [\underline{0} \vdots \underline{V}_2(t)] \qquad \underline{V}_2(t) \varepsilon M_{n(m-m_1)} \quad . \tag{4.3.4}$$

All such $\underline{V}(t)$ are of rank $\leq m - m_1 \triangleq m_2$.

Theorem 4.3.1: Let $\underline{V}(t)$ be of the form (4.3.4); if there exists an observer $O_T^{2c}(L)$, $\underline{T}(t) \varepsilon \mathcal{J}_V^c$, which is compatible, then rank $\underline{V}(t) = m_2$ and rank $(\underline{I}_n - \underline{V}(t)\underline{C}(t)) = n - m_2$.

Proof: By lemma 4.2.3, we may assume that there exists $\Theta_T^{1c}(L) \varepsilon \pi_V^{1c}(L)$ which is compatible. Let $\mathcal{E}_T^{2c}(L)$ be the associated estimator and $\underline{w}(t)$ the resulting estimate. Using (4.3.4), we have

$$\underline{e}(t) = \underline{w}(t)-\underline{x}(t) = \underline{P}(t)\underline{z}(t)+\underline{V}(t)\underline{y}(t)-\underline{x}(t)$$

$$= \underline{P}(t)\underline{z}(t)+\underline{V}_2(t)\underline{C}_2(t)\underline{x}(t)-\underline{x}(t) = \underline{P}(t)(\underline{z}(t)-\underline{T}(t)\underline{x}(t)) \quad . \tag{4.3.5}$$

† Compatibility is defined as in discrete case. See section 3.4.

By compatibility we must have

$$\underline{C}_2(t)\underline{e}(t) = \underline{C}_2(t)\underline{P}(t)(\underline{z}(t) - \underline{T}(t)\underline{x}(t)) = \underline{0} \quad \text{a.s.} \tag{4.3.6}$$

Thus in particular if $\underline{\xi}(t) \equiv \underline{0}$, $\underline{\eta}(t) \equiv \underline{0}$, (4.3.6) implies that

$$\underline{C}_2(t)\underline{P}(t)\underline{\Phi}_F(t,t_c^+)\underline{T}(t_o^+)\underline{\alpha} = \underline{0} \quad ; \quad \underline{\alpha} \in R^n \text{ arbitrary} \tag{4.3.7}$$

where $\underline{F}(t)$ is given by (4.2.9). Since $\Theta_T^{2c}(L) \in \pi_V^{pc}(L)$, therefore $\underline{T}(t_c^+)$ may be assumed to be of full rank, and so (4.3.7) implies that

$$\underline{C}_2(t)\underline{P}(t) = \underline{0} \quad . \tag{4.3.8}$$

Using (4.3.5), we have

$$\underline{C}_2(t)\underline{e}(t) = \underline{C}_2(t)\underline{V}_2(t)\underline{C}_2(t)\underline{x}(t) - \underline{C}_2(t)\underline{x}(t) = \underline{0} \quad \text{a.s.} \tag{4.3.9}$$

$\underline{x}(t)$ can be an arbitrary vector in R^n; so we conclude that

$$\underline{C}_2(t)\underline{V}_2(t) = \underline{I}_{m_2} \tag{4.3.10}$$

and that rank $\underline{V}_2(t) = m_2$.

From (4.3.8), we have

$$\underline{C}_2(t)\underline{P}(t)\underline{T}(t) = \underline{0} \quad ; \quad \text{rank } \underline{P}(t)\underline{T}(t) \leq n - m_2 \tag{4.3.11}$$

$\underline{P}(t)$, $\underline{T}(t)$ satisfy

$$\underline{P}(t)\underline{T}(t) + \underline{V}_2(t)\underline{C}_2(t) = \underline{I}_n \quad . \tag{4.3.12}$$

Equations (4.3.12) and (4.3.10) imply rank $\underline{P}(t)\underline{T}(t) \geq n - m_2$. Together with (4.3.11) we have

$$\text{rank}\left(\underline{I}_n - \underline{V}(t)\underline{C}(t)\right) = \text{rank } \underline{P}(t)\underline{T}(t) = n - m_2 \quad . \tag{4.3.13}$$

By theorems 4.2.4 and 4.3.1, we see that all $\underline{V}(t)$ of the form (4.3.4) can be classified into two classes: either all $\mathcal{O}_T^{2c}(L)$, $\underline{T}(t) \in \mathfrak{J}_V^c$ are all uncompatible or all $\mathcal{O}_T^{2c}(L)$, $\underline{T}(t) \in \mathfrak{J}_V^c$ are all compatible and equivalent. We may call the former class of $\underline{V}(t)$ uncompatible and the latter class of $\underline{V}(t)$ compatible. Thus, the classes of observers which we shall consider are parameterized by $\underline{L}(t) \in M_{nm}$ and compatible $\underline{V}(t)$ of the form (4.3.4).

Let $\underline{V}(t)$ be a fixed, differentiable matrix function of the form (4.3.4) which is compatible; and let $\underline{L}(t) \in M_{nm}$ be arbitrary. From theorem 4.2.4, all observers $\mathcal{O}_T^{2c}(L)$, $\underline{T}(t) \in \mathfrak{J}_V^c$ are equivalent and thus yield the same error dynamics. Let $\mathcal{O}_T^{2c}(L) \in \mathfrak{m}_V^{2c}(L)$, its associated estimator $\mathcal{E}_T^{2c}(L)$ is described by:

$$\mathcal{E}_T^{2c}(L): \begin{aligned} \dot{\underline{z}}(t) &= (\underline{T}(t)(\underline{A}(t)-\underline{L}(t)\underline{C}(t))\underline{P}(t)+\dot{\underline{T}}(t)\underline{P}(t))\underline{z}(t)+(\underline{T}(t)(\underline{A}(t)-\underline{L}(t)\underline{C}(t))\underline{V}_2(t) \\ &\quad +\dot{\underline{T}}(t)\underline{V}_2(t)+\underline{T}(t)\underline{L}_2(t))\underline{y}_2(t)+\underline{T}(t)\underline{L}_1(t)\underline{y}_1(t)+\underline{T}(t)\underline{B}(t)\underline{u}(t) \\ \underline{w}(t) &= \underline{P}(t)\underline{z}(t)+\underline{V}_2(t)\underline{y}_2(t) \end{aligned} \tag{4.3.14}$$

where

$$\underline{L}(t) = [\underline{L}_1(t) \vdots \underline{L}_2(t)] \quad ; \quad \underline{L}_1(t) \in M_{nm_1} \quad , \quad \underline{L}_2(t) \in M_{nm_2} \tag{4.3.15}$$

$\underline{P}(t) \in M_{n(n-m_2)}$, $\underline{T}(t) \in M_{(n-m_2)n}$ satisfy (4.3.12) and, in addition, they satisfy:

$$\underline{T}(t)\underline{V}_2(t) = \underline{0}_{nm_2} \quad ; \quad \underline{C}_2(t)\underline{P}(t) = \underline{0}_{m_2 n} \quad ; \quad \underline{C}_2(t)\underline{V}_2(t) = \underline{I}_{m_2} \quad ;$$

$$\underline{T}(t)\underline{P}(t) = \underline{I}_{m_2} \quad . \tag{4.3.16}$$

We can simplify the structure of $\mathcal{E}_T^{2c}(L)$ by using (4.3.15) and (4.3.16):

$$\underline{T}(t)\underline{L}(t)\underline{C}(t)\underline{P}(t) = \underline{T}(t)(\underline{L}_1(t)\underline{C}_1(t)+\underline{L}_2(t)\underline{C}_2(t))\underline{P}(t) = \underline{T}(t)\underline{L}_1(t)\underline{C}_1(t)\underline{P}(t) \tag{4.3.17}$$

$$\underline{T}(t)\underline{L}_2(t) - \underline{T}(t)\underline{L}(t)\underline{C}(t)\underline{V}_2(t) = -\underline{T}(t)\underline{L}_1(t)\underline{C}_1(t)\underline{V}_2(t) \tag{4.3.18}$$

Substituting (4.3.17), (4.3.18) into (4.3.14), the structure of the estimator $\mathcal{E}_T^{2c}(L)$ is given by (Figure 4.3)

$$\mathcal{E}_T^{2c}(L): \quad \begin{aligned} \dot{\underline{z}}(t) &= (\underline{T}(t)\underline{A}(t)\underline{P}(t)+\dot{\underline{T}}(t)\underline{P}(t)-\underline{T}(t)\underline{L}_1(t)\underline{C}_1(t)\underline{P}(t))\underline{z}(t)+\underline{T}(t)\underline{B}(t)\underline{u}(t) \\ &\quad +\underline{T}(t)\underline{L}_1(t)\underline{y}_1(t)+(\underline{T}(t)\underline{A}(t)\underline{V}_2(t)+\dot{\underline{T}}(t)\underline{V}_2(t)-\underline{T}(t)\underline{L}_1(t)\underline{C}_1(t)\underline{V}_2(t))\underline{y}_2(t) \\ \underline{w}(t) &= \underline{P}(t)\underline{z}(t)+\underline{V}_2(t)\underline{y}_2(t) \end{aligned} \tag{4.3.19}$$

By demanding $\mathcal{E}_T^{2c}(L)$ to give unbiased estimates of $\underline{x}(t)$, we set (see also section 4.5)

$$\underline{z}(t_o) = \underline{T}(t_o)\underline{x}_o \tag{4.3.20}$$

where $\underline{T}(t_o)$, $\underline{P}(t_o)$ satisfy

$$\underline{P}(t_o)\underline{T}(t_o) + \underline{V}_2(t_o)\underline{C}_2(t_o) = \underline{I}_n \quad . \tag{4.3.21}$$

Using (4.3.1), (4.3.2), (4.3.16), (4.3.19) and (4.3.20), we have the dynamics of the error process, $\underline{e}(t) \doteq \underline{w}(t) - \underline{x}(t)$, given by: (see Appendix C)

$$\begin{aligned} \dot{\underline{e}}(t) &= (\underline{A}(t)-\underline{V}_2(t)\bar{\underline{C}}_2(t)-\underline{P}(t)\underline{T}(t)\underline{L}_1(t)\underline{C}_1(t))\underline{e}(t)+(\underline{V}_2(t)\underline{C}_2(t)-\underline{I}_n)\underline{\xi}(t) \\ &\quad +\underline{P}(t)\underline{T}(t)\underline{L}_1(t)\underline{\theta}(t) \\ \underline{e}(t_o) &= (\underline{I}_n - \underline{V}_2(t_c)\underline{C}_2(t_o))(\underline{x}_o - \underline{x}(t_o)) \end{aligned} \tag{4.3.22}$$

and

$$\bar{\underline{C}}_2(t) \doteq \dot{\underline{C}}_2(t) + \underline{C}_2(t)\underline{A}(t) \quad . \tag{4.3.23}$$

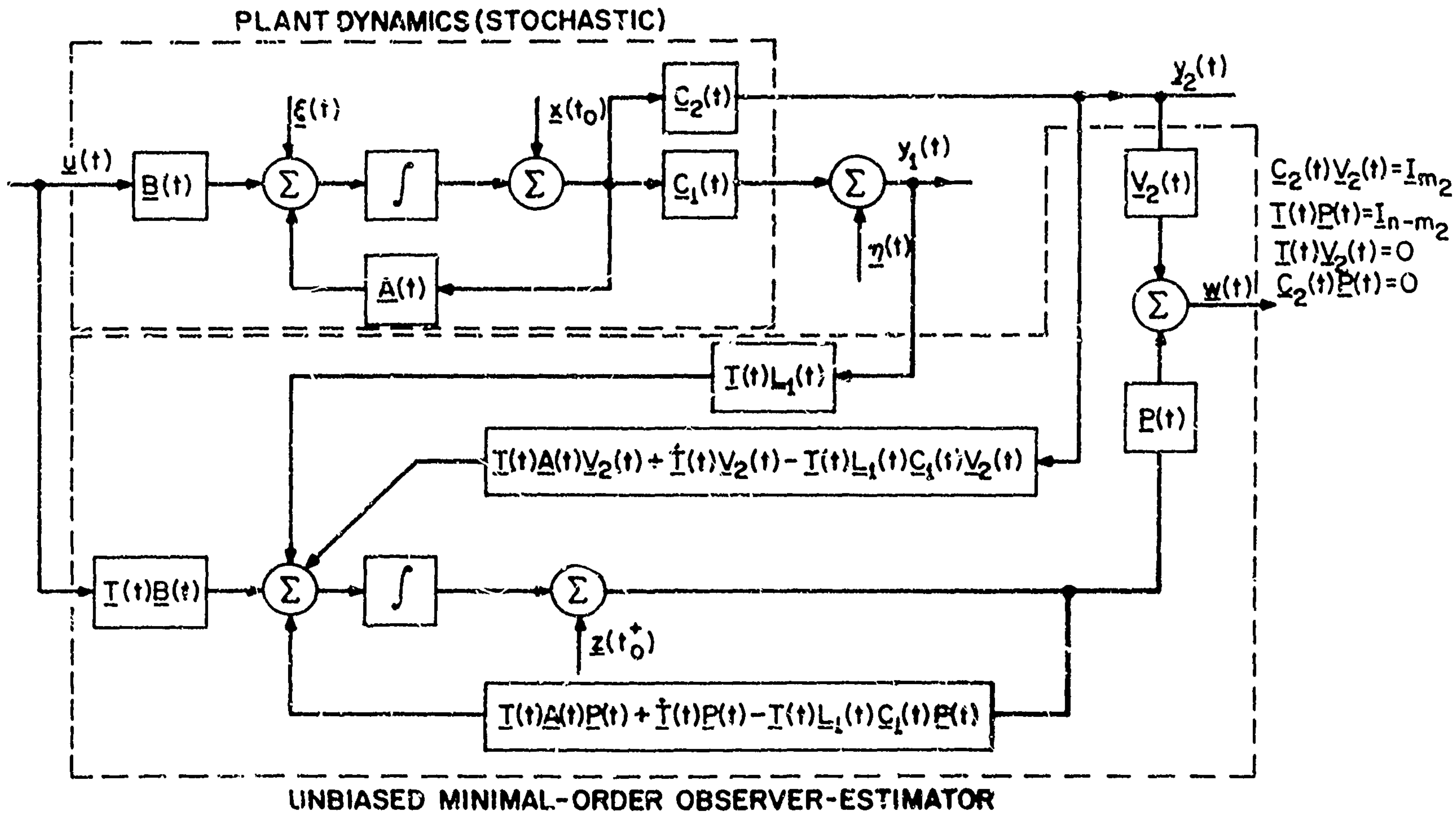


Fig. 4.3 THE STRUCTURE OF AN UNBIASED MINIMAL-ORDER OBSERVER-ESTIMATOR

Define

$$\underline{\tilde{L}}_1(t) \triangleq \underline{P}(t)\underline{T}(t)\underline{L}_1(t) \ \varepsilon \ M_{nm_2} \quad . \qquad (4.3.24)$$

The error covariance is given by (see chapter 2, section 2.3)

$$\begin{aligned}\dot{\underline{\Sigma}}(t) = &[\underline{A}(t)-\underline{V}_2(t)\underline{\tilde{C}}_2(t)-\underline{\tilde{L}}_1(t)\underline{C}_1(t)]\underline{\Sigma}(t)+\underline{\Sigma}(t)[\underline{A}(\ \ -\underline{V}_2(t)\underline{\dot{C}}_2(t)-\underline{\tilde{L}}_1(t)\underline{C}_1(t)]' \\ &+[\underline{I}_n-\underline{V}_2(t)\underline{C}_2(t)]\underline{R}(t)[\underline{I}_n-\underline{V}_2(t)\underline{C}_2(t)]'+\underline{\tilde{L}}_1(t)\underline{Q}(t)\underline{\tilde{L}}_1'(t)\end{aligned}$$

$$\underline{\Sigma}(t_o) = (\underline{I}_n - \underline{V}_2(t_o)\underline{C}(t_o))\underline{\Sigma}_o(\underline{I}_n - \underline{V}_2(t_o)\underline{C}_2(t_o))' \qquad (4.3.25)$$

We note that the dynamics of the error covariance are dependent on $\underline{V}_2(t)$, $t \in [t_c,\infty]$ and $\underline{\tilde{L}}_1(t)$, $t \in [t_o,\infty]$. Note that $\underline{V}_2(t)$ and $\underline{\tilde{L}}_1(t)$ are not arbitrary but $\underline{V}_2(t)$ has to satisfy (4.3.10) and $\underline{\tilde{L}}_1(t)$ is related to $\underline{L}_1(t)$, which is an arbitrary matrix, by equation (4.3.24). To find the optimum class of observers, we are to find a pair $\{\underline{V}_2^*(t), \underline{\tilde{L}}_1^*(t)\}$, which may be a nonunique pair, with the above constraints which will give the "least" nonnegative definite covariance matrix. Each such pair $\{\underline{V}_2^*(t), \underline{\tilde{L}}_1^*(t)\}$ specifies an optimum class of observers. When $\{\underline{V}_2(t), \underline{\tilde{L}}_1(t)\}$ ranges over all constrained pairs, we generate the solution set, $\tilde{\beta}_{t_o}$, of (4.3.25). The minimization problem is equivalent to finding $\{\underline{V}_2^*(t), \underline{\tilde{L}}_1^*(t)\}$ which will yield the minimal function with respect to the solution set $\tilde{\beta}_{t_o}$.

Theorem 4.3.2: Let $\underline{C}_2(t)\underline{R}(t)\underline{C}_2'(t) > \underline{0}$, then there exists a unique constrained pair $\{\underline{V}_2^*(t),\ t \in [t_o,\infty],\ \underline{\tilde{L}}_1^*(t),\ t \in [t_o,\infty]\}$ which yield the unique minimal function, $\underline{\Sigma}^*(t)$, with respect to $\tilde{\beta}_{t_o}$. $\underline{V}_2^*(t)$, $\underline{\tilde{L}}_1^*(t)$ are given by

$$\underline{V}_2^*(t) = \begin{cases} \underline{\Sigma}_o\underline{C}_2'(t_o)(\underline{C}_2(t_o)\underline{\Sigma}_o\underline{C}_2'(t_o))^{-1} & t = t_o \\ (\underline{\Sigma}^*(t)\underline{\tilde{C}}_2'(t) + \underline{R}(t)\underline{C}_2'(t))\underline{\Lambda}^{-1}(t) & t > t_o \end{cases} \qquad (4.3.26)$$

$$\underline{\tilde{L}}_1^*(t) = \underline{\Sigma}^*(t)\underline{C}_1'(t)\underline{Q}^{-1}(t) \qquad t > t_o \tag{4.3.27}$$

where

$$\underline{\Delta}(t) \triangleq \underline{C}_2(t)\underline{R}(t)\underline{C}_2'(t) > \underline{0} \tag{4.3.28}$$

$\underline{\Sigma}^*(t)$ is the minimal function with respect to $\tilde{\mathcal{B}}_{t_o}$ and is given by

$$\begin{aligned}\dot{\underline{\Sigma}}^*(t) = {} & (\underline{A}(t)-\underline{R}(t)\underline{C}_2'(t)\underline{\Delta}^{-1}(t)\underline{\tilde{C}}_2(t))\underline{\Sigma}^*(t)+\underline{\Sigma}^*(t)(\underline{A}(t)-\underline{R}(t)\underline{C}_2'(t)\underline{\Delta}^{-1}(t)\underline{\tilde{C}}_2(t))' \\ & -\underline{\Sigma}^*(t)(\underline{\tilde{C}}_2'(t)\underline{\Delta}^{-1}(t)\underline{\tilde{C}}_2(t)+\underline{C}_1'(t)\underline{Q}^{-1}(t)\underline{C}_1(t))\underline{\Sigma}^*(t)+\underline{R}(t) \\ & -\underline{R}(t)\underline{C}_2'(t)\underline{\Delta}^{-1}(t)\underline{C}_2(t)\underline{R}(t)\end{aligned}$$

$$\underline{\Sigma}^*(t_o) = \underline{\Sigma}_o - \underline{\Sigma}_o\underline{C}_2'(t_o)(\underline{C}_2(t_o)\underline{\Sigma}_o\underline{C}_2'(t_o))^{-1}\underline{C}_2(t_o)\underline{\Sigma}_o \tag{4.3.29}$$

Let $\mathcal{B}_{t_o}$ be the solution set of (4.3.25) when $\{\underline{V}_2(t), \underline{\tilde{L}}_1(t)\}$ ranges <u>all possible pairs; $\underline{\Sigma}^*(t)$ is also the minimal function with respect to $\mathcal{B}_{t_o}$; thus $\underline{\Sigma}^*(t)$ is the Riccati function (see definition 2.6.4).</u>

<u>Proof</u>: Let $\mathcal{B}_{t_o}$ be the solution set of (4.3.25) when $\{\underline{V}_2(t), \underline{\tilde{L}}_1(t)\}$ ranges all possible pairs. Compare (4.3.25) and (2.6.1) with

$$\begin{aligned}
\underline{\Sigma}^*(t) &\rightarrow \underline{P}_V(t,t_o;\underline{\Sigma}^*(t_o)) \\
[\underline{\tilde{L}}_1(t) \vdots \underline{V}_2(t)] &\rightarrow \underline{V}(t) \\
\begin{bmatrix}\underline{C}_1(t)\\ \cdots \\ \underline{\tilde{C}}_2(t)\end{bmatrix} &\rightarrow \underline{D}_1(t) \\
\begin{bmatrix}\underline{Q}(t) & \vdots & \underline{0}\\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{0}\end{bmatrix} &\rightarrow \underline{Q}(t) \\
\begin{bmatrix}\underline{0}\\ \cdots \\ \underline{\tilde{C}}_2(t)\end{bmatrix} &\rightarrow \underline{D}_2(t)
\end{aligned} \tag{4.3.30}$$

Since by assumption $\underline{C}_2(t)\underline{R}(t)\underline{C}_2'(t) > \underline{0}$, then we have

$$\begin{bmatrix} \underline{Q}(t) & \vdots & \underline{0} \\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{0} \end{bmatrix} + \begin{bmatrix} \underline{0} \\ \cdots \\ \underline{C}_2(t) \end{bmatrix} \underline{R}(t)[\underline{0} \; \vdots \; \underline{C}_2'(t)] = \begin{bmatrix} \underline{Q}(t) & \vdots & \underline{0} \\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{C}_2(t)\underline{R}(t)\underline{C}_2'(t) \end{bmatrix} > \underline{0} \quad (4.3.31)$$

and so the unique minimal function is given by: (see (2.6.19) to (2.6.21))

$$\dot{\underline{\Sigma}}^*(t) = (\underline{A}(t)-\underline{R}(t)\underline{C}_2'(t)\underline{\Delta}^{-1}(t)\tilde{\underline{C}}_2(t))\underline{\Sigma}^*(t)+\underline{\Sigma}^*(t)(\underline{A}(t)-\underline{R}(t)\underline{C}_2'(t)\underline{\Delta}^{-1}(t)\tilde{\underline{C}}_2(t))'$$
$$-\underline{\Sigma}^*(t)(\tilde{\underline{C}}_2(t)\underline{\Delta}^{-1}(t)\tilde{\underline{C}}_2(t)+\underline{C}_1'(t)\underline{Q}^{-1}(t)\underline{C}_1(t))\underline{\Sigma}^*(t)+R(t)-\underline{R}(t)\underline{C}_2'(t)\underline{\Delta}^{-1}(t)\underline{C}_2(t)\underline{R}(t)$$

$$\underline{\Sigma}^*(t_o) = \underline{\Sigma}_o - \underline{\Sigma}_o\underline{C}_2'(t_o)(\underline{C}_2(t_o)\underline{\Sigma}_o\underline{C}_2'(t_o))^{-1}\underline{C}_2(t_o)\underline{\Sigma}_o$$

(4.3.32)

and $[\tilde{\underline{L}}_1^*(t) \; \vdots \; \underline{V}_2^*(t)]$ is given by: (see 2.6.18)

$$[\tilde{\underline{L}}_1^*(t) \; \vdots \; \underline{V}_2^*(t)] = \underline{\Sigma}^*(t)[\underline{C}_1'(t) \; \vdots \; \underline{C}_2(t)]+[\underline{0} \; \vdots \; \underline{R}(t)\underline{C}_2'(t)]\begin{bmatrix} \underline{Q}^{-1}(t) & \vdots & \underline{0} \\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{\Delta}^{-1}(t) \end{bmatrix}$$

$$t > t_o$$

$$\underline{V}_2^*(\;_o) = \underline{\Sigma}_o\underline{C}_2'(t_o)(\underline{C}_2(t_o)\underline{\Sigma}_o\underline{C}_2'(t_o))^{-1}$$

(4.3.33)

To complete the prove of the theorem, we need only to show that $\underline{V}_2^*(t)$ satisfies (4.3.10), and $\tilde{\underline{L}}_1^*(t)$ is related to some $\underline{L}_1^*(t) \in M_{nm_1}$ via (4.3.24).

From (4.3.32), we see that

$$\underline{C}_2(t_o)\underline{\Sigma}^*(t_o) = \underline{0} \quad (4.3.34)$$

Using (4.3.32), (4.3.28) and (4.3.23) we have

$$\frac{d}{dt}(\underline{C}_2(t)\underline{\Sigma}^*(t)) = \underline{C}_2(t)\underline{\Sigma}^*(t)[(\underline{A}(t)-\underline{R}(t)\underline{C}_2'(t)\underline{\Delta}^{-1}(t)\tilde{\underline{C}}_2(t))'-(\tilde{\underline{C}}_2(t)\underline{\Delta}^{-1}(t)\tilde{\underline{C}}_2(t)$$
$$+\underline{C}_1'(t)\underline{Q}^{-1}(t)\underline{C}_1(t))\underline{\Sigma}^*(t)] \quad (4.3.35)$$

Thus we conclude that

$$\underline{C}_2(t)\underline{\Sigma}^*(t) = \underline{0} \tag{4.3.36}$$

Therefore we have

$$\underline{C}_2(t)\underline{V}_2^*(t) = [\underline{C}_2(t)\underline{\Sigma}^*(t)\underline{C}_1'(t) + \underline{\Delta}(t)]\underline{\Delta}^{-1}(t) = \underline{I}_{m_2} \quad . \tag{4.3.37}$$

Let

$$\underline{L}_1^*(t) = \underline{\Sigma}^*(t)\underline{C}_1'(t)\underline{Q}^{-1}(t) = \tilde{\underline{L}}_1^*(t) \quad . \tag{4.3.38}$$

We can easily see from (4.3.30) that

$$\begin{aligned}\underline{T}(t)\underline{P}(t)\underline{L}_1^*(t) &= (\underline{I}_n-\underline{V}_2(t)\underline{C}_2(t))\underline{\Sigma}^*(t)\underline{C}_1'(t)\underline{Q}^{-1}(t) = \underline{\Sigma}^*(t)\underline{C}_1'(t)\underline{Q}^{-1}(t)\\ &= \tilde{\underline{L}}_1^*(t) \quad .\end{aligned} \tag{4.3.39}$$

Thus $\underline{\Sigma}^*(t)$, given by (4.3.31), is also the minimal function with respect to $\tilde{\beta}_{t_o}$.

We now have the structure of a class of minimal order optimum observers, $0_T^{2c}(L^*)$, $\underline{T}(t) \varepsilon \mathfrak{J}_{V^*}^c$, ($\underline{V}^*(t) = [\underline{0} \,\vdots\, \underline{V}_2^*(t)]$) and its associated estimators $\mathcal{E}_T^{2c}(L^*)$, $\underline{T}(t) \varepsilon \mathfrak{J}_{V^*}^c$:

$$\mathcal{E}_T^{2c}(L^*): \quad \begin{aligned}\dot{\underline{z}}^*(t) = {} & (\underline{T}(t)\underline{A}(t)\underline{P}(t)+\dot{\underline{T}}(t)\underline{P}(t)-\underline{T}(t)\underline{L}_1^*(t)\underline{C}_1(t)\underline{P}(t))\underline{z}^*(t)+\underline{T}(t)\underline{L}_1^*(t)\underline{y}_1(t)\\ & +(\underline{T}(t)\underline{A}(t)\underline{V}_2^*(t)+\dot{\underline{T}}(t)\underline{V}_2^*(t)-\underline{T}(t)\underline{L}_1^*(t)\underline{C}_1(t)\underline{V}_2^*(t)\underline{y}_2(t)\\ & +\underline{T}(t)\underline{B}(t)\underline{u}(t)\end{aligned}$$

$$\underline{w}^*(t) = \underline{P}(t)\underline{z}^*(t) + \underline{V}_2^*(t)\underline{y}_2(t) \; ; \; \underline{z}^*(t_o = \underline{T}(t_o)\underline{x}_o \tag{4.3.40}$$

with $\underline{V}_2^*(t)$, $\underline{L}_1^*(t)$ given by (4.3.32), (4.3.33) and (4.3.37); $\underline{P}(t)$, $\underline{T}(t)$ satisfy

$$\underline{T}(t)\underline{V}_2^*(t) = \underline{0}_{nm_2} \quad ; \quad \underline{C}_2(t)\underline{P}(t) = \underline{0}_{m_2 n} \quad ; \quad \underline{T}(t)\underline{P}(t) = \underline{I}_{m_2} \quad ; \quad t \geq t_o \tag{4.3.41}$$

$\underline{C}_2(t)$ is continuous $t \in [t_o,\infty]$, thus we can choose $\underline{P}(t)$ which is continuous for all $t \in [t_o,\infty]$. From theorem 4.3.2, $\underline{V}_2^*(t)$ is discontinuous at $t = t_o$, and so from (4.3.41), $\underline{T}(t)$ is discontinuous at t_o. Since $\underline{e}^*(t) \triangleq \underline{w}^*(t) - \underline{x}(t)$ is continuous at $t = t_o$, we have

$$\underline{e}^*(t_o) = \underline{P}(t_o)(\underline{z}^*(t_o)-\underline{T}(t_o)\underline{x}(t_o)) = \underline{P}(t_o)(\underline{z}^*(t_o^+)-\underline{T}(t_o^+)\underline{x}(t_o)) \tag{4.3.42}$$

and using (4.3.40) and the fact that $\underline{z}^*(t_o) = \underline{T}(t_o)\underline{x}_o$, we have

$$\underline{z}^*(t_o^+) = \underline{T}(t_o)\underline{x}_o - \underline{T}(t_o)\underline{V}_2^*(t_o^+)\underline{y}_2(t_o) \quad . \tag{4.3.43}$$

We see that $\underline{z}^*(t)$ is discontinuous at $t = t_o$, and consists of the <u>a priori</u> guess $(\underline{T}(t_o)\underline{x}_o)$ and a correction term due to perfect observation $(\underline{T}(t_o)\underline{V}_2^*(t_o^+)\underline{y}_2(t_o^+))$. The detail structure of $\mathcal{E}_T^{2c}(L^*)$, $\underline{T}(t) \in \mathcal{J}_{V^*}^c$, is shown in Figure 4.4. What we have obtained is <u>a class of optimum mean square estimators among a restricted class of estimators being considered.</u> For example, we have not considered the class of nonlinear estimators. Now to prove the derived minimum order optimum observer-estimator is the truly optimum estimator, we appeal to the projection theorem. It is clear that $\underline{w}^*(t)$ is a linear functional of $\underline{y}_2(s)$, $s \in [t_o,t]$ and $\underline{y}_1(s)$, $s \in [t_o,t]$, we shall prove that the error process, $\underline{e}^*(t) \triangleq \underline{w}^*(t) - \underline{x}(t)$, satisfies the projection equations

$$E\{\underline{e}^*(t)\underline{y}_1'(s)\} = \underline{0} \quad , \quad s \in [t_o,t) \quad ; \quad E\{\underline{e}^*(t)\underline{y}_2'(t)\} = \underline{0} \quad , \quad s \in [t_o,t] \tag{4.3.44}$$

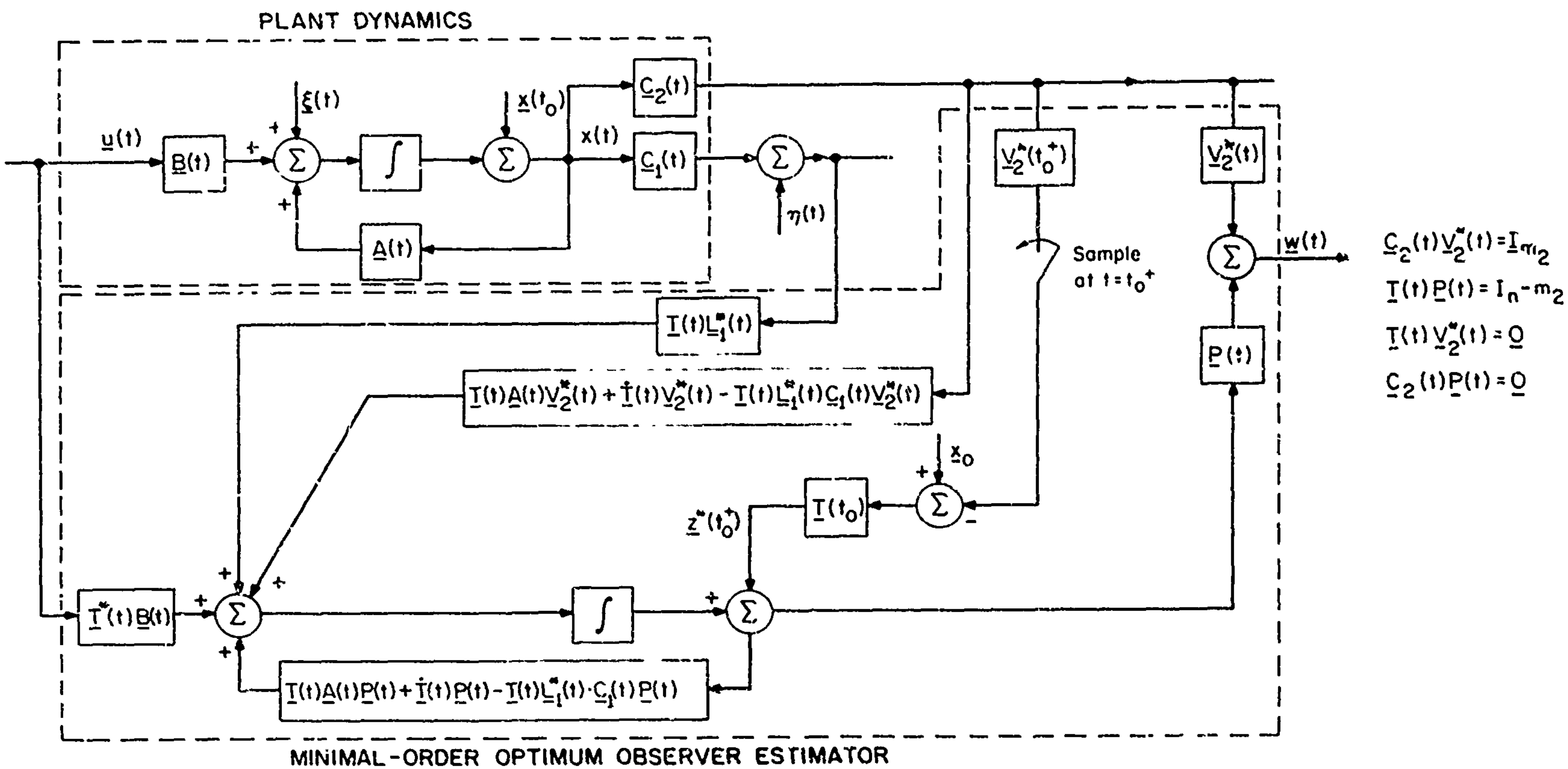


Fig. 4.4 THE STRUCTURE OF MINIMAL-ORDER OPTIMUM OBSERVER-ESTIMATOR

This implies that (see discrete analog and Appendix B) the optimum class of observers will yield (a.s.) the conditional mean estimates of $\underline{x}(t)$, and thus reveals the truly optimum nature of $\mathcal{E}_T^{2c}(L^*)$, $\underline{T}(t) \in \mathcal{J}_V^c$.

By using $\mathcal{E}_T^{2c}(L^*)$, $\underline{T}(t) \in \mathcal{J}_{V^*}^c$, as an estimating device, the corresponding error process will satisfy

$$\dot{\underline{e}}^*(t) = (\underline{A}(t)-\underline{V}_2^*(t)\tilde{\underline{C}}_2(t)-\underline{L}_1^*(t)\underline{C}_1(t))\underline{e}^*(t)+(\underline{V}_2^*(t)\underline{C}_2(t)-\underline{I}_n)\underline{\xi}(t)+\underline{L}_1^*(t)\underline{n}(t)$$

$$\underline{e}^*(t_o) = (\underline{I}_n - \underline{V}_2^*(t_o)\underline{C}_2(t_o))(\underline{x}_o - \underline{x}(t_o)) \tag{4.3.45}$$

where $\underline{V}_2^*(t)$, $\underline{L}_1^*(t)$ are given by (4.3.32), (4.3.33) and (4.3.37).

<u>Lemma 4.3.3</u>: Let $\{\underline{e}^*(t),\ t \geq t_o\}$ be a random process satisfying (4.3.45); and $\underline{x}(t)$, $t \geq t_o$, be described by (4.3.1) with $\underline{u}(t) \equiv \underline{0}$, $t \geq t_o$. Then for all $t \geq t_o$, we have

$$E\{\underline{e}(t)\underline{x}'(t)\} = -\underline{\Sigma}^*(t) \quad . \tag{4.3.46}$$

<u>Proof</u>: At $t = t_o$, we have from (4.3.44), (4.3.33) and (4.3.31) that

$$E\{\underline{e}^*(t_o)\underline{x}'(t_o)\} = \{\underline{I}_n - \underline{\Sigma}_o\underline{C}_2'(t_o)(\underline{C}_2(t_o)\underline{\Sigma}_o\underline{C}_2'(t_o))^{-1}\underline{C}_2(t_o)\}\{\underline{x}_o\underline{x}_o'-E\ \underline{x}(t_o)\underline{x}'(t_o)\}$$

$$= -\underline{\Sigma}^*(t_o) \quad . \tag{4.3.47}$$

Using (4.3.45) and (4.3.1) we have

$$\begin{aligned}
\frac{dE\{\underline{e}^*(t)\underline{x}'(t)\}}{dt} &= (\underline{A}(t)-\underline{V}_2^*(t)\tilde{\underline{C}}_2(t)-\underline{L}_1^*(t)\underline{C}_1(t))E\{\underline{e}^*(t)\underline{x}'(t)\} \\
&\quad +(\underline{V}_2^*(t)\underline{C}_2(t)-\underline{I}_n)\underline{R}(t)+E\{\underline{e}^*(t)\underline{x}'(t)\}\underline{A}'(t) \\
&= (\underline{A}(t)-\underline{R}(t)\underline{C}_2'(t)\underline{\Delta}^{-1}(t)\tilde{\underline{C}}_2(t))E\{\underline{e}^*(t)\underline{x}'(t)\} \\
&\quad -\underline{\Sigma}^*(t)\underline{C}_1'(t)\underline{Q}^{-1}(t)\underline{C}_1(t)E\{\underline{e}^*(t)\underline{x}'(t)\} \\
&\quad -\underline{\Sigma}^*(t)\tilde{\underline{C}}_2'(t)\underline{\Delta}^{-1}(t)\tilde{\underline{C}}_2(t)E\{\underline{e}^*(t)\underline{x}'(t)\}+E\{\underline{e}^*(t)\underline{x}'(t)\}\underline{A}'(t) \\
&\quad +\underline{\Sigma}^*(t)\tilde{\underline{C}}_2'(t)\underline{\Delta}^{-1}(t)\underline{C}_2(t)\underline{R}(t)-\underline{R}(t)+\underline{R}(t)\underline{C}_2'(t)\underline{\Delta}^{-1}(t)\underline{C}_2(t)\underline{R}(t)
\end{aligned} \tag{4.3.48}$$

Let us define

$$\underline{D}(t) = \underline{\Sigma}^*(t) + E\{\underline{e}^*(t)\underline{x}'(t)\} \quad . \tag{4.3.49}$$

By (4.3.31), (4.3.47) and (4.3.48) we have

$$\dot{\underline{D}}(t) = (\underline{A}(t)-\underline{R}(t)\underline{C}_2'(t)\underline{\Lambda}^{-1}(t)\tilde{\underline{C}}_2(t))\underline{D}(t)-\underline{\Sigma}^*(t)(\underline{C}_1'(t)\underline{Q}^{-1}(t)\underline{C}_1(t)+\tilde{\underline{C}}_2'(t)\underline{\Lambda}^{-1}(t)\tilde{\underline{C}}_2(t))\cdot$$
$$\underline{D}(t)+\underline{D}'(t)\underline{A}(t)$$

$$\underline{D}(t_o) = \underline{0} \tag{4.3.50}$$

$\underline{\Sigma}^*(t)$ is the unique minimal sequence of $\mathcal{B}_{t_o}$ and ε is well defined. (4.3.50) implies

$$\underline{D}(t) = \underline{0} \;\; ; \;\; t \geq t_o \tag{4.3.51}$$

and (4.3.46) follows.

<u>Theorem 4.3.4</u>: Let $\underline{e}^*(t)$, $t \geq t_o$, be described by (4.3.45), and $\underline{y}(t)$, $t \geq t_o$, be given by (4.3.1) with $\underline{u}(t) \equiv \underline{0}$, $t \geq t_o$. Then for all $t \geq t_o$:

$$E\{\underline{e}^*(t)\underline{y}_1'(s)\} = \underline{0} \quad , \quad s \in [t_o,t) \quad ; \quad E\{\underline{e}^*(t)\underline{y}_2'(s)\} = \underline{0} \quad , \quad s \in [t_o,t] \tag{4.3.52}$$

<u>Proof</u>: By (4.3.45) and (4.3.1) and the properties of Gaussian white noise:

$$E\{\underline{e}^*(t)\underline{y}_1'(s)\} = E\{\underline{e}^*(t)\underline{x}'(s)\}\underline{C}_1'(s)+E\{\underline{e}^*(t)\underline{n}'(s)\}$$
$$= \underline{\Phi}_{\tilde{A}}(t,s)[E\{\underline{e}^*(s)\underline{x}'(s)\}\underline{C}_1'(s)+\underline{L}_1^*(s)\underline{Q}(s)] \qquad s \in [t_o,t) \tag{4.3.53}$$

where

$$\tilde{\underline{A}}(t) = \underline{A}(t) - \underline{V}_2^*(t)\tilde{\underline{C}}_2(t) - \underline{L}_1^*(t)\underline{C}_1(t) \tag{4.3.54}$$

and $\underline{\Phi}_{\hat{A}}(t,s)$ is the fundamental matrix associated with $\hat{\underline{A}}(t)$. Now using lemma 4.3.3, (4.3.53) and (4.3.27) imply

$$E\{\underline{e}^*(t)\underline{v}_1'(s)\} = \underline{\Phi}_{\hat{A}}(t,s)[-\underline{\Sigma}^*(s)\underline{C}_1'(s) + \underline{\Sigma}^*(s)\underline{C}_1'(s)] = \underline{0} \quad . \quad (4.3.55)$$

Similarly for $s \in [t_0,t]$, we have (by using compatibility)

$$E\{\underline{e}^*(t)\underline{v}_2'(s)\} = \underline{\Phi}_{\hat{A}}(t,s)E\{\underline{e}^*(t)\underline{x}'(s)\}\underline{C}_2'(s) = -\underline{\Phi}_{\hat{A}}(t,s)\underline{\Sigma}^*(s)\underline{C}_2'(s) = \underline{0} \quad . \quad (4.3.56)$$

The above theorem implies that for zero control, the optimum class of observers and their associated estimators will all generate (a.s.) the conditional mean estimates of $\underline{x}(t)$. The results also holds if $\underline{u}(t)$ is a nonzero but known deterministic control function, because we can always subtract its deterministic contribution. The case where the control is generated via a special class of feedback law will be considered in chapter V. Note that we obtain the Kalman Filter as a special case when we set $\underline{C}_2(t) = \underline{0}$ (4.3.29).

4.4 Asymptotic Behavior of Estimators

Let us first consider the asymptotic behavior of classes of observers and associated estimators for a deterministic system $\mathcal{S}_1^c$. Then, we shall consider the asymptotic behavior of optimum classes of observers and associated estimator for the stochastic system $\mathcal{S}_2^c$.

Definition 4.4.1: The system $\mathcal{S}_1^c$ is detectable at τ if there exists an observer $\mathcal{O}_T^{1c}(L)$, $\underline{T}(t) \in \mathcal{J}_V^c$, and its associated estimator $\mathcal{E}_T^{1c}(L)$:

$$\mathcal{E}_T^{1c}(L): \begin{aligned} \dot{\underline{z}}(t) &= (\underline{T}(t)(\underline{A}(t)-\underline{L}(t)\underline{C}(t))\underline{P}(t)+\dot{\underline{T}}(t)\underline{P}(t))\underline{z}(t)+[\underline{T}(t)(\underline{A}(t)-\underline{L}(t)\underline{C}(t))\underline{V}(t) \\ &\quad +\dot{\underline{T}}(t)\underline{V}(t)+\underline{T}(t)\underline{L}(t)]\underline{y}(t)+\underline{T}(t)\underline{B}(t)\underline{u}(t) \end{aligned}$$

$$\underline{w}(t) = \underline{P}(t)\underline{z}(t)+\underline{V}(t)\underline{y}(t) \quad ; \quad \underline{z}(\tau) \in S_\tau = \{\underline{T}(\tau)\underline{\alpha}|\underline{\alpha} \in R^n\} \qquad (4.4.1)$$

such that for all $\underline{z}(\tau) \varepsilon S_\tau$, $\underline{w}(t) \to \underline{x}(t)$, as $t \to \infty$. The system $\mathcal{S}_1^c$ is said to be detectable if it is detectable at $\tau \varepsilon (-\infty,\infty)$.

We shall say a stochastic system $\mathcal{S}_2^c$ to be detectable if its deterministic analog, $\mathcal{S}_1^c$, is detectable.

<u>Theorem 4.4.2</u>: The system $\mathcal{S}_1^c$ is detectable if and only if there exists an observer, $\mathcal{O}_T^{1c}(L)$, $\mathcal{J}_V^c$, which is uniformly asymptotically stable.

<u>Proof</u>: The estimation error by using any observer, $\mathcal{O}_T^{1c}(L)$, $\underline{T}(t) \varepsilon \mathcal{J}_V^c$, and its associated estimator is given by (see equation (4.3.5))

$$\underline{e}(t) = \underline{P}(t)\{\underline{z}(t) - \underline{T}(t)\underline{x}(t)\} \triangleq \underline{P}(t)\underline{\tilde{z}}(t,\tau;\underline{\tilde{z}}_o) \qquad (4.4.2)$$

where $\underline{\tilde{z}}(t,\tau;\underline{\tilde{z}}_o)$ satisfies (see equation (4.2.6) and theorem 4.2.2)

$$\dot{\underline{\tilde{z}}}(t,\tau;\underline{\tilde{z}}_o) = [\underline{T}(t)(\underline{A}(t)-\underline{L}(t)\underline{C}(t))\underline{P}(t)+\dot{\underline{T}}(t)\underline{P}(t)]\underline{\tilde{z}}(t,\tau;\underline{\tilde{z}}_o) \quad ;$$
$$\underline{\tilde{z}}(\tau,\tau;\underline{\tilde{z}}_o) = \underline{\tilde{z}}_o \ \varepsilon \ S_\tau \quad . \qquad (4.4.3)$$

Let us first assume that there exists some $\underline{L}(t) \varepsilon M_{nm}$ and $\underline{V}(t) \varepsilon M_{nm}$ such that an observer $\mathcal{O}_T^{1c}(L)$, $\underline{T}(t) \varepsilon \mathcal{J}_V^c$, is uniformly asymptotically stable; then for all τ and $\underline{\tilde{z}}_o \varepsilon S_\tau$, $\underline{\tilde{z}}(t,\tau;\underline{\tilde{z}}_o) \to \underline{0}$ as $t \to \infty$. From (4.4.2) we conclude that $\mathcal{S}_1^c$ is detectable. Conversely, if the system $\mathcal{S}_1^c$ is detectable, then there exists an observer, $\mathcal{O}_T^{1c}(\tilde{L})$, $\underline{T}(t) \varepsilon T_{\tilde{V}}^c$, such that the output of its associated estimator will give exact asymptotic estimates independent of when we initiate the observer state; i.e., for all τ, and $\underline{\tilde{z}}_o \varepsilon S_\tau$

$$\underline{e}(t) = \underline{P}(t)\underline{\tilde{z}}(t,\tau;\underline{\tilde{z}}_o) \to \underline{0} \quad \text{as } t \to \infty \qquad (4.4.4)$$

where $\underline{\tilde{z}}(t,\tau;\underline{\tilde{z}}_o)$ is given by (4.4.3). We may assume $\underline{P}(t)$ to be of full rank; thus (4.4.4) implies that the system (4.4.3) is uniformly asymptotically

stable, and so the observer $\mathcal{O}_{\underline{T}}^{1c}(\tilde{L})$, $\underline{T}(t) \in \mathcal{J}_V^c$, is uniformly asymptotically stable.

A linear system

$$\dot{\underline{x}}(t) = \underline{A}(t)\underline{x}(t) \tag{4.4.5}$$

is said to be exponentially stable if there exists α_1, $\alpha_2 > 0$ such that

$$||\underline{\Phi}_A(t,\tau)|| \le \alpha_1 e^{-\alpha_2|t-\tau|} \tag{4.4.6}$$

where $\underline{\Phi}_A(t,\tau)$ is the fundamental matrix of $\underline{A}(t)$. We also say that the matrix $\underline{A}(t)$ is exponentially stable. Theorem 4.4.2 relates detectability to the structure of the observers. Since exponentially stable implies uniform asymptotic stability and vise versa [48], the above lemma implies that the system $\mathcal{S}_1^c$ is detectable if and only if there exists an observer $\mathcal{O}_{\underline{T}}^{1c}(L)$, $\underline{T}(t) \in \mathcal{J}_V^c$, such that the error of estimates (in the noise free case) by using $\mathcal{E}_{\underline{T}}^{1c}(L)$, $\underline{T}(t) \in \mathcal{J}_V^c$, has the bound

$$||e(t)|| \le \alpha_1' e^{-\alpha_2|t-t_o|} \tag{4.4.7}$$

where t_o is the initial time. We may call such an estimator $\mathcal{E}_{\underline{T}}^{1c}(L)$ an exponential estimator.[49]

Theorem 4.4.3: If there exists $\underline{V}_2(t) \in M_{nm_2}$ of rank m_2 and $\underline{\tilde{L}}_1(t) \in M_{nm_1}$ such that $(\underline{A}(t) - \underline{V}_2(t)\underline{\tilde{C}}_2(t) - \underline{\tilde{L}}_1(t)\underline{C}_1(t))$ is exponentially stable, then the equivalent classes of observers $\mathcal{O}_{\underline{T}}^{1c}(\tilde{L})$, $\underline{T}(t) \in \mathcal{J}_V^c$, where $\underline{V}(t) = [\underline{0} \vdots \underline{V}_2(t)]$ and $\underline{L}(t) = [\underline{\tilde{L}}_1(t) \vdots \underline{\tilde{L}}_2(t)]$, $\underline{\tilde{L}}_2(t)$ arbitrary, are all uniformly asymptotically stable and so $\mathcal{E}_{\underline{T}}^{1c}(\tilde{L})$, $\underline{T}(t) \in \mathcal{J}_V^c$, will yield exponentially consistent estimates.

Proof: Let us consider the class of $\underline{V}(t)$ of the form (4.3.4) with rank $\underline{V}_2(t) = n_2$. The error, $\underline{e}(t) \triangleq \underline{w}(t) - \underline{x}(t)$, of estimates by using $\underline{V}(t)$ within this class is given by (see also (4.3.22))

$$\dot{\underline{e}}(t) = (\underline{A}(t) - \underline{V}_2(t)\tilde{\underline{C}}_2(t) - \tilde{\underline{L}}_1(t)\underline{C}_1(t))\underline{e}(t) \quad . \tag{4.4.8}$$

By assumption, there exists $\underline{V}_2(t) \in M_{nm_2}$ of rank n_2 and $\tilde{\underline{L}}_1(t) \in M_{nm_2}$, such that $(\underline{A}(t) - \underline{V}_2(t)\tilde{\underline{C}}_2(t) - \tilde{\underline{L}}_1(t)\underline{C}_1(t))$ is exponentially stable, thus the theorem follows from (4.4.8) and theorem 4.4.2.

Theorem 4.4.3 gives us a sufficiency test for detectability; it also indicates how we can construct an exponential estimator.

For the stochastic system S_2^c, the class of minimal order optimum observers and their associated estimators are given by (4.3.39) and (4.3.42); the optimum error covariance $\underline{\Sigma}^*(t)$ is given by (4.3.29). We shall now investigate the asymptotic behavior of the class of minimal order optimum estimators via $\underline{\Sigma}^*(t)$.

Theorem 4.4.4: The matrix function $\underline{\Sigma}^*(t)$ will remain bounded for all $t \in [t_o,\infty]$ if and only if there exists $\underline{V}_2(t) \in M_{nm_2}$ and $\tilde{\underline{L}}_1(t) \in M_{nm_1}$, such that $(\underline{A}(t) - \underline{V}_2(t)\tilde{\underline{C}}_2(t) - \tilde{\underline{L}}_1(t)\underline{C}_1(t))$ is exponentially stable.

Proof: This follows immediately from theorem 4.3.2. The reader is referred to the proof in the discrete analog for the detailed argument. Using theorem 4.4.2 and 4.4.3, we see that $\Theta_T^{2c}(L^*)$, $\underline{T}(t) \in \mathcal{J}_{V^*}^c$, is uniformly asymptotically stable.

Corollary 4.4.5: If $(\underline{A}(t),\underline{C}(t))$ is uniformly completely observable, i.e., there exists $\infty > \tau > 0$ such that

$$\underline{M}(\tau) = \int_t^{t+\tau} \underline{\phi}_A'(\sigma,t+\tau)\underline{C}'(\sigma)\underline{C}(\sigma)\underline{\phi}_A(\sigma,t+\tau)d\sigma \quad ; \quad t \in [t_o,\infty] \tag{4.4.9}$$

has rank n, then there exists $\underline{L}(t) \in M_{nm}$ such that $(\underline{A}(t) - \underline{L}(t)\underline{C}(t))$ is exponentially stable.

Proof: Apply theorem 4.4.4 to the special case when $m_2 = 0$; i.e., all observation channels are corrupted by white Gaussian noise. The optimum error covariance will remain bounded if and only if there exists $\underline{L}(t) \in M_{nm}$ such that $(\underline{A}(t) - \underline{L}(t)\underline{C}(t))$ is exponentially stable. If $(\underline{A}(t),\underline{C}(t))$ is uniformly completely observable, the optimum error covariance will remain bounded for all $t \in [t_o,\infty]$,[50] and so the corollary follows.

Let us consider the time invariant case where $\underline{A}$, $\underline{C}_1$, and $\underline{C}_2$ are constant and bounded matrices.

Lemma 4.4.6: If the pair $(\underline{A},\underline{C})$ is observable, then the pair

$$\left(\underline{A},\begin{bmatrix}\underline{C}_1\\ \cdots \\ \tilde{\underline{C}}_2\end{bmatrix}\right)$$

is also observable.

Proof: Construct the matrix

$$\tilde{\underline{M}}(t_o,\tau) = \int_{t_o}^{\tau} \underline{\phi}_A'(\sigma,\tau)[\underline{C}_1' \ \vdots\ \tilde{\underline{C}}_2']\begin{bmatrix}\underline{C}_1\\ \cdots \\ \tilde{\underline{C}}_2\end{bmatrix}\underline{\phi}_A(\sigma,\tau)d\sigma$$

$$= \int_{t_o}^{\tau} \underline{\phi}_A'(\sigma,\tau)\underline{C}_1'\underline{C}_1\underline{\phi}_A(\sigma,\tau)d\sigma + \int_{t_o}^{\tau} \frac{d\underline{\phi}_A'(\sigma,\tau)\underline{C}'_2}{d\sigma} \cdot \frac{d\underline{C}_2\underline{\phi}_A(\sigma,\tau)}{d\sigma}\, d\sigma \quad . \tag{4.4.10}$$

Let $\underline{x} \in R^n$ such that $\underline{x}'\tilde{\underline{M}}(t_o,\tau)\underline{x} = 0$; then from (4.4.10) we have for $\sigma \in [t_o,\tau]$:

$$\underline{C}_1\underline{\phi}_A(\sigma,\tau)\underline{x} = \underline{0}_{m_2} \quad ; \quad \underline{C}_2\underline{\phi}_A(\sigma,\tau)\underline{x} = \underline{y} \in R^m \tag{4.4.11}$$

where $\underline{y}$ is a constant vector. Suppose that $\underline{x} \neq \underline{0}$; let $\underline{x}_o = \underline{\phi}_A(t_o,\tau)\underline{x}$; then $\underline{x}_o \neq \underline{0}$; let $\underline{x}_1 = \underline{\phi}_A(t,\tau)\underline{x}$, $\tau > t_1 > t_o$, also $\underline{x}_1 \neq \underline{0}$. Since $\underline{A}$ is constant,

we have from (4.4.11) that

$$\underline{C}_2\underline{\phi}_A(\sigma,t_o)\underline{x}_o = \underline{C}_2\underline{\phi}_A(\sigma,\tau)\underline{x} = \underline{y} \; ; \quad \sigma \in [t_o,\tau] \tag{4.4.12}$$

$$\begin{aligned}\underline{C}_2\underline{\phi}_A(\sigma,t_o)\underline{x}_1 &= \underline{C}_2\underline{\phi}_A(\sigma,t_o)\underline{\phi}_A(t_1,\tau)\underline{x} = \underline{C}_2\underline{\phi}_A(\sigma + t_1 - t_o,t_1)\underline{\phi}_A(t_1,\tau)\underline{x} \\ &= \underline{C}_2\underline{\phi}_A(\sigma + t_1 - t_o,\tau)\underline{x} = \underline{y} \; ; \quad \sigma \in [t_o,\tau - t_1 + t_o] \quad .\end{aligned} \tag{4.4.13}$$

Thus $\underline{x}_o$, $\underline{x}_1$ are indistinguishable by observing the output in the nonzero interval $[t_o,\tau - t_1 + t_o]$. This contradicts the assumption that $(\underline{A},\underline{C})$ is observable.

<u>Lemma 4.4.7</u>: Let $\underline{\Sigma}_o = \underline{0}$; the solution of (4.3.29), denoted by $\underline{\Sigma}^*(t;\underline{0})$ will reach a steady state $\underline{\Sigma}^*$ which satisfies

$$\begin{aligned}\underline{0} = {} & (\underline{A}-\underline{R}\,\underline{C}_2'\underline{\Delta}^{-1}\underline{C}_2\underline{A})\underline{\Sigma}^* + \underline{\Sigma}^*(\underline{A}-\underline{R}\,\underline{C}_2'\underline{\Delta}^{-1}\underline{C}_2\underline{A})' - \underline{\Sigma}^*(\underline{A}'\underline{C}_2'\underline{\Delta}^{-1}\underline{C}_2\underline{A}+\underline{C}_1'\underline{Q}^{-1}\underline{C}_1)\underline{\Sigma}^* \\ & + \underline{R}-\underline{R}\,\underline{C}_2'\underline{\Delta}^{-1}\underline{C}_2\underline{R} \quad ; \qquad \underline{\Delta} = \underline{C}_2\underline{R}\,\underline{C}_2' > \underline{0}\end{aligned} \tag{4.4.14}$$

if and only if there exists $\underline{V}_2(t)$, $\underline{\tilde{L}}_1(t)$ such that $(\underline{A} - \underline{V}_2(t)\underline{\tilde{C}}_2 - \underline{\tilde{L}}_1(t)\underline{C}_1)$ is exponentially stable.

<u>Proof</u>: Let us consider $\underline{\Sigma}^*(t;\underline{0})$ as a minimal function with respect to the solution set $\mathcal{B}_{t_o}$. With the assumption that $\underline{\Sigma}_o = \underline{0}$, we have $\underline{\Sigma}^*(t_o;\underline{0}) = \underline{0}$ from (4.3.25), and so

$$\underline{\Sigma}^*(t;\underline{0}) = \int_{t_o}^{t} \underline{\phi}(t,\tau)\{(\underline{I}_n-\underline{V}_2^*(\tau)\underline{C}_2)\underline{R}(\underline{I}_n-\underline{V}_2^*(\tau)\underline{C}_2)'+\underline{\tilde{L}}_1^*(\tau)\underline{Q}\,\underline{\tilde{L}}_1^{*\prime}(\tau)\}\underline{\phi}'(t,\tau)d\tau \tag{4.4.15}$$

where $\underline{\phi}(t,\tau)$ is the fundamental matrix associated with $(\underline{A} - \underline{V}_2^*(t)\underline{\tilde{C}}_2 - \underline{\tilde{L}}_1^*(t)\underline{C}_1)$, and $\underline{V}_2^*(t)$, $\underline{\tilde{L}}_1^*(t)$ are given by

$$\underline{V}_2^*(t) = (\underline{\Sigma}^*(t;\underline{0})\underline{\tilde{C}}_2' + \underline{R}\ \underline{C}_2')\underline{\Delta}^{-1} \quad ; \quad t > t_o \tag{4.4.16}$$

$$\underline{L}_1^*(t) = \underline{\Sigma}^*(t;\underline{0})\underline{C}_1'\ \underline{Q}^{-1} \quad ; \quad t > t_o \quad . \tag{4.4.17}$$

Let $\underline{V}_2^\sigma(t) = \underline{V}_2^*(t + \sigma)$, $\underline{\tilde{L}}_1^\sigma(t) = \underline{\tilde{L}}_1^*(t + \sigma)$, and $\underline{\phi}^\sigma(t,\tau)$ be the fundamental matrix associated with $(\underline{A} - \underline{V}_2^\sigma(t)\underline{\tilde{C}}_2 - \underline{\tilde{L}}_1^\sigma(t)\underline{C}_1)$. Clearly we have

$$\underline{\phi}^\sigma(t,\tau) = \underline{\phi}(t + \sigma,\tau + \sigma) \quad . \tag{4.4.18}$$

Let $\underline{\Sigma}^\sigma(t;\underline{0})$, $t \geq t_o$ be the solution of (4.3.25) with $\underline{\Sigma}^\sigma(t_o;\underline{0}) = \underline{0}$, thus $\underline{\Sigma}^\sigma(t;\underline{0}) \in \mathcal{B}_{t_o}$. Since $\underline{\Sigma}^*(t)$ is the minimal function with respect to $\mathcal{B}_{t_o}$, we have

$$\underline{\Sigma}^*(t;\underline{0}) \leq \underline{\Sigma}^\sigma(t;\underline{0}) \quad ; \quad t \geq t_o \quad . \tag{4.4.19}$$

Also we have from (4.4.18) and the definition of $\underline{V}_2^\sigma(t)$, $\underline{\tilde{L}}_1^\sigma(t)$ that:

$$\underline{\Sigma}^\sigma(t-\sigma;\underline{0}) = \int_{t_o}^{t-\sigma} \underline{\phi}^\sigma(t-\sigma,\tau)\{(\underline{I}_n-\underline{V}_2^\sigma(\tau)\underline{C}_2)\underline{R}(\underline{I}_n-\underline{V}_2^\sigma(\tau)\underline{C}_2)+\underline{\tilde{L}}_1^\sigma(\tau)\underline{Q}\ \underline{\tilde{L}}_1^{\sigma\prime}(\tau)\}\underline{\phi}^{\sigma\prime}(t-\sigma,\tau)d\tau$$

$$\leq \int_{t_o}^{t} \underline{\phi}(t,\gamma)\{(\underline{I}_n-\underline{V}_2^*(\gamma)\underline{C}_2)\underline{R}(\underline{I}_n-\underline{V}_2^*(\gamma)\underline{C}_2)+\underline{\tilde{L}}_1^*(\gamma)\underline{Q}\ \underline{\tilde{L}}_1^*(\gamma)\}\underline{\phi}'(t,\gamma)d\gamma$$

$$= \underline{\Sigma}^*(t) \quad . \tag{4.4.20}$$

Combining (4.4.19) and (4.4.20) we have

$$\underline{\Sigma}^*(t;\underline{0}) \geq \underline{\Sigma}^\sigma(t - \sigma;\underline{0}) \geq \underline{\Sigma}^*(t - \sigma;\underline{0}) \quad . \tag{4.4.21}$$

The lemma follows from theorem 4.4.4 and the monotone nondecreasing nature of $\underline{\Sigma}^*(t;\underline{0})$ as t increases (4.4.21).

Theorem 4.4.8: For all $\underline{\Sigma}_o > \underline{0}$, the solution of (4.3.29), denoted by $\underline{\Sigma}^*(t;\underline{\Sigma}_o)$ will reach a steady state $\underline{\Sigma}^*$ which satisfies (4.4.14) if and only

there exists $\underline{V}_2(t)$, $\tilde{\underline{L}}_1(t)$ such that $(\underline{A} - \underline{V}_2(t)\tilde{\underline{C}}_2 - \tilde{\underline{L}}_1(t)\underline{C}_1)$ is exponentially stable.

<u>Proof</u>: From (4.3.29) and (4.3.25) we have

$$\underline{\Sigma}^*(t_o;\underline{\Sigma}_o) - \underline{\Sigma}^*(t_o;\underline{0}) = \underline{\Sigma}_o - \underline{\Sigma}_o\underline{C}_2'(\underline{C}_2\underline{\Sigma}_o\underline{C}_2')^{-1}\underline{C}_2\underline{\Sigma}_o \geq \underline{0} \quad . \tag{4.4.22}$$

Therefore, from (2.6.21) we deduce that

$$\underline{\Sigma}^*(t;\underline{\Sigma}_o) \geq \underline{\Sigma}^*(t;\underline{0}) \ , \quad t \in [t_o,\infty] \quad . \tag{4.4.23}$$

Using the minimal property of $\underline{\Sigma}^*(t;\underline{\Sigma}_o)$, we have

$$\underline{0} \leq \underline{\Sigma}^*(t;\underline{\Sigma}_o) - \underline{\Sigma}^*(t;\underline{0}) \leq \underline{\phi}(t,t_o)\underline{\Sigma}^*(t_o;\underline{\Sigma}_o)\underline{\phi}'(t,t_o) \tag{4.4.24}$$

where $\underline{\phi}(t,t_o)$ is the fundamental matrix associated with $(\underline{A} - \underline{V}_2^*(t)\tilde{\underline{C}}_2 - \tilde{\underline{L}}_1(t)\underline{C}_1)$ and $\underline{V}_2^*(t)$, $\tilde{\underline{L}}_1^*(t)$ are given by (4.4.16) and (4.4.17). $\underline{\phi}(t,t_o)$ is exponentially stable if and only if there exists $\underline{V}_2(t)$, $\tilde{\underline{L}}_1(t)$ such that $(\underline{A} - \underline{V}_2(t)\tilde{\underline{C}}_2 - \tilde{\underline{L}}_1(t)\underline{C}_1)$ is exponentially stable. Using lemma 4.4.7, we obtain the theorem easily.

From lemma 4.4.6, and corollary 4.4.5, we see that observability of the pair $(\underline{A},\underline{C})$ is <u>sufficient</u> to assure that $\underline{\Sigma}^*(t;\underline{\Sigma}_o) \to \underline{\Sigma}^*$ satisfying (4.4.14) where $\underline{\Sigma}_o > \underline{0}$ is arbitrary.

4.5 <u>General Discussion</u>

In this chapter, we considered the estimation of deterministic and stochastic systems using the observer approach.

In the deterministic case, sufficient conditions for existence of exponential estimator have been derived; such estimators can be realized by an observer $\mathcal{O}_T^{1c}(L)$, $\underline{T}(t) \in \mathcal{J}_V^c$, which is asymptotically stable, and its associated estimator $\mathcal{E}_T^{1c}(L)$.

In the stochastic case, the minimal order optimum observer and its estimator are described in detail in Figure 4.4. The optimum error covariance, $\underline{\Sigma}^*(t)$, is given by (4.3.29). Asymptotic behavior of the minimal order optimum observer is investigated via the optimum error covariance $\underline{\Sigma}^*(t)$. Necessary and sufficient condition for $\underline{\Sigma}^*(t)$ to be uniformly bounded have been established. The condition is related closely to the structural property of the system $\mathcal{S}_2^c$ under consideration.

In the following, we shall discuss different points which are relevant to the whole development in this chapter.

(A) <u>Unbiased Estimates and Observer-Estimator Structure</u>

Let $\mathcal{S}_2^c$ be a stochastic system described by (4.3.1) with $\underline{u}(t) \equiv \underline{0}$. Let an unbiased estimator $\mathcal{E}$ be given by

$$\underline{w}(t) = \int_{t_o}^{t} \underline{H}(t,\tau)\underline{y}(\tau)d\tau + \underline{V}(t)\underline{y}(t) \tag{4.5.1}$$

where $\underline{H}(\cdot,\cdot)$ is an nxm matrix whose elements are differentiable in both arguments. Since $E\{\underline{w}(t)\} = E\{\underline{x}(t)\}$, from (4.3.1) and (4.5.1) we have

$$\int_{t_o}^{t} \underline{H}(t,\tau)\underline{C}(\tau)\underline{\phi}_A(\tau,t_o)\underline{x}_o d\tau + \underline{V}(t)\underline{C}(t)\underline{\phi}_A(t,t_o)\underline{x}_o = \underline{\phi}_A(t,t_o)\underline{x}_o \tag{4.5.2}$$

where $\underline{\phi}_A(t,\tau)$ is the fundamental matrix associated with $\underline{A}(t)$. The structure of the estimator should be independent of the mean of $\underline{x}(t_o)$, $\underline{x}_o$, thus (4.5.2) implies

$$\int_{t_o}^{t} \underline{H}(t,\tau)\underline{C}(\tau)\underline{\phi}_A(\tau,t_o)d\tau + \underline{V}(t)\underline{C}(t)\underline{\phi}_A(t,t_o) = \underline{\phi}_A(t,t_o) \quad . \tag{4.5.3}$$

Differentiate both sides of (4.5.3) in respect to t.

$$\underline{H}(t,t)\underline{C}(t)+\int_{t_o}^{t} \frac{\partial \underline{H}(t,\tau)}{\partial t} \underline{C}(\tau)\underline{\phi}_A(\tau,t)d\tau+\underline{\dot{V}}(t)\underline{C}(t)+\underline{V}(t)\underline{\dot{C}}(t)+\underline{V}(t)\underline{C}(t)\underline{A}(t) = \underline{A}(t) \quad . \tag{4.5.4}$$

Multiplying both sides of (4.5.4) by $\underline{w}(t)$ and taking expectations

$$\int_{t_o}^{t}\left\{(\underline{H}(t,t)\underline{C}(t)+\dot{\underline{V}}(t)\underline{C}(t)+\underline{V}(t)\dot{\underline{C}}(t)+\underline{V}(t)\underline{C}(t)\underline{A}(t)-\underline{A}(t))\underline{H}(t,\tau)+\frac{\partial H(t,\tau)}{\partial t}\right\}E\{\underline{y}(\tau)\}d\tau$$

$$+(\underline{H}(t,t)\underline{C}(t)+\dot{\underline{V}}(t)\underline{C}(t)+\underline{V}(t)\dot{\underline{C}}(t)+\underline{V}(t)\underline{C}(t)-\underline{A}(t))\underline{V}(t)E\{\underline{y}(\tau)\}=\underline{\mathbf{0}} \qquad (4.5.5)$$

(4.5.5) is satisfied if $\underline{H}(t,\tau)$ and $\underline{V}(t)$ satisfy

$$\int_{t_o}^{t}\left(\underline{G}(t)\underline{H}(t,\tau)+\frac{\partial H(t,\tau)}{\partial t}\right)\underline{y}(\tau)d\tau = -\underline{G}(t)\underline{V}(t)\underline{y}(t) \qquad (4.5.6)$$

where $\underline{y}(t)$ is $\underline{a}$ m-vector valued function of t; and

$$\underline{G}(t) = \underline{H}(t,t)\underline{C}(t) + \dot{\underline{V}}(t)\underline{C}(t) + \underline{V}(t)\dot{\underline{C}}(t) + \underline{V}(t)\underline{C}(t)\underline{A}(t) - \underline{A}(t) \quad . \qquad (4.5.7)$$

Let us denote $\underline{w}_1(t) = \int_{t_o}^{t} \underline{H}(t,\tau)\underline{y}(\tau)d\tau$; we have

$$\dot{\underline{w}}_1(t) = \underline{H}(t,t)\underline{y}(t) + \int_{t_o}^{t} \frac{\partial H(t,\tau)}{\partial t}\underline{y}(\tau)d\tau$$

$$= \underline{H}(t,t)\underline{y}(t) - \underline{G}(t)\underline{w}_1(t) - \underline{G}(t)\underline{V}(t)\underline{y}(t) \quad . \qquad (4.5.8)$$

The unbiased estimator is realized by

$$\mathcal{E}: \quad \begin{aligned} \dot{\underline{w}}(t) &= -\underline{G}(t)\underline{w}_1(t) + (\underline{H}(t,t) - \underline{G}(t)\underline{V}(t))\underline{y}(t) \\ \underline{w}(t) &= \underline{w}_1(t) + \underline{V}(t)\underline{y}(t) \quad . \end{aligned} \qquad (4.5.9)$$

By some transformation of coordinates, the unbiased estimator $\mathcal{E}$ can be realized by

$$\mathcal{E}': \quad \begin{aligned} \dot{\underline{z}}(t) &= \underline{F}(t)\underline{z}(t) + \underline{D}(t)\underline{y}(t) \\ \underline{w}(t) &= \underline{P}(t)\underline{z}(t) + \underline{V}(t)\underline{y}(t) \quad . \end{aligned} \qquad (4.5.10)$$

Since $\mathcal{E}'$ an unbiased estimator, we have at $t = t_o$

$$\underline{x}_o = \underline{P}(t_o)\underline{z}(t_o) + \underline{V}(t_o)\underline{C}(t_o)\underline{x}_o \quad . \qquad (4.5.11)$$

If $\underline{T}(t_o) \in \mathcal{I}(\underline{C}(t_o);m,s,n)$ and $\underline{P}(t_o)$ is such that

$$\underline{P}(t_o)\underline{T}(t_o) + \underline{V}(t_o)\underline{C}(t_o) = \underline{I}_n \quad , \qquad (4.5.12)$$

then by setting $\underline{z}(t_o) = \underline{T}(t_o)\underline{x}_o$, (4.5.11) is clearly satisfied. Also we have $E\{\underline{w}(t)\} = E\{\underline{x}(t)\}$ thus

$$\underline{P}(t)\ \underline{\Phi}_F(t,t_o)\underline{T}(t_o)\underline{x}_o + \int_{t_o}^{t} \underline{\Phi}_F(t,\tau)\underline{D}(\tau)\underline{C}(\tau)\underline{\Phi}_A(\tau,t_o)\underline{x}_o d\tau + \underline{V}(t)\underline{C}(t)\underline{\Phi}_A(t,t_o)\underline{x}_o$$

$$= \underline{\Phi}_A(t,t_o)\underline{x}_o \quad . \qquad (4.5.13)$$

The structure of the estimator is required to be independent of $\underline{x}_o$, therefore (4.5.13) implies

$$\underline{P}(t)\underline{T}(t) + \underline{V}(t)\underline{C}(t) = \underline{I}_n \qquad (4.5.14)$$

where $\underline{T}(t)$ is given by

$$\underline{T}(t) = \underline{\Phi}_F(t,t_o)\underline{T}(t_o)\underline{\Phi}_A(t_o,t) + \int_{t_o}^{t} \underline{\Phi}_F(t,\tau)\underline{D}(\tau)\underline{C}(\tau)\underline{\Phi}_A(\tau,t)d\tau \quad . \qquad (4.5.15)$$

By comparing with (4.2.4) and theorem 4.2.2, such an estimator can be realized by an observer $O_T^{1c}(L)$, $\underline{T}(t) \in \mathcal{S}_V^c$ and its associated estimator $\mathcal{E}_T^{1c}(L)$, where $\underline{L}(t) \in M_{nm}$ is arbitrary.

Thus we see that the concept of an observer is in some sense equivalent to the concept of unbiased estimator. When the a priori distribution of $\underline{x}(t_o)$ is known, the minimal order optimum observer-filter gives unbiased minimum mean square estimates; whereas if the _a priori_ distribution

of $\underline{x}(t_o)$ is unknown, the minimal order optimum observer-filter will be an asymptotically unbiased minimum mean square estimator.

(B) Estimation for Linear System

The observer theorem introduced in this chapter generalizes and unifies estimation theory for deterministic and stochastic systems. For both deterministic and stochastic cases, the structure of the estimators are the same. In the deterministic case, we are to find certain parameters, $\underline{V}_2(t)$, $\tilde{\underline{L}}_1(t)$, so as to obtain exponentially consistent estimates, whereas in the stochastic case, the optimum choice of $\underline{V}_2(t)$ and $\tilde{\underline{L}}_1(t)$ is specified by the noise statistical law and the detailed structure of the system. Thus we see that in the deterministic case, qualitative theory should be used in designing well-behaved observer-estimator;[49] whereas in the stochastic case, optimization technique can be applied to derive the class of minimum order optimum observer-estimator.

(C) Kalman Filtering Technique

We can also solve the stochastic problem in section 4.3 by using the Kalman Filtering Approach.† Let us consider the system S_2^c with $\underline{u}(t) \equiv \underline{0}$.

Let $\underline{T}(t) \in M_{n(n-m_1)}$ such that

$$\begin{bmatrix} \underline{T}(t) \\ \cdots \\ \underline{C}_2(t) \end{bmatrix}$$

is of full rank. Define

$$\underline{x}_1(t) = \underline{T}(t)\underline{x}(t) \quad . \qquad (4.5.16)$$

†This approach was suggested by I. B. Rhodes.

Then we have

$$\underline{x}(t) = \underline{P}(t)\underline{x}_1(t) + \underline{V}_2(t)\underline{y}_2(t) \tag{4.5.17}$$

where

$$[\underline{P}(t) \vdots \underline{V}_2(t)] = \begin{bmatrix} \underline{T}(t) \\ \cdots \\ \underline{C}_2(t) \end{bmatrix}^{-1} \quad . \tag{4.5.18}$$

We have the equation for $\underline{x}_1(t)$ and $\underline{y}_1(t)$

$$\begin{aligned} \dot{\underline{x}}_1(t) &= (\dot{\underline{T}}(t)+\underline{T}(t)\underline{A}(t))\underline{x}(t)+\underline{T}(t)\underline{\xi}(t) \\ &= (\dot{\underline{T}}(t)\underline{P}(t)+\underline{T}(t)\underline{A}(t)\underline{P}(t))\underline{x}_1(t)+(\dot{\underline{T}}(t)\underline{V}_2(t)+\underline{T}(t)\underline{A}(t)\underline{V}_2(t))\cdot \\ &\quad \underline{y}_2(t)+\underline{T}(t)\underline{\xi}(t) \\ &\quad \underline{y}_1(t)-\underline{C}_1(t)\underline{V}_2(t)\underline{y}_2(t) = \underline{C}_1(t)\underline{P}(t)\underline{x}_1(t)+\underline{\eta}(t) \end{aligned} \tag{4.5.19}$$

Since $\underline{y}_2(t)$ can be observed exactly, we can assume it is known. Now apply a Kalman filter to the system (4.5.19): the best mean square estimate of $\underline{x}_1(t)$ is given by

$$\begin{aligned} \dot{\hat{\underline{x}}}_1(t) &= (\dot{\underline{T}}(t)\underline{P}(t)+\underline{T}(t)\underline{A}(t)\underline{P}(t)-\hat{\underline{\Sigma}}(t)\underline{P}'(t)\underline{C}_1(t)\underline{Q}^{-1}(t)\underline{C}_1(t)\underline{P}(t))\hat{\underline{x}}_1(t) \\ &\quad +(\dot{\underline{T}}(t)\underline{V}_2(t)+\underline{T}(t)\underline{A}(t)\underline{V}_2(t))\underline{y}_2(t) \\ &\quad +\hat{\underline{\Sigma}}(t)\underline{P}'(t)\underline{C}_1(t)\underline{Q}^{-1}(t)(\underline{y}_1(t)-\underline{C}_1(t)\underline{V}_2(t)\underline{y}_2(t)) \end{aligned} \tag{4.5.20}$$

$$\hat{\underline{x}}_1(t_o) = \underline{T}(t_o)\underline{x}_o$$

and $\hat{\underline{\Sigma}}(t)$ satisfies

$$\begin{aligned} \dot{\hat{\underline{\Sigma}}}(t) &= (\dot{\underline{T}}(t)\underline{P}(t)+\underline{T}(t)\underline{A}(t)\underline{P}(t)\hat{\underline{\Sigma}}(t)+\hat{\underline{\Sigma}}(t)(\dot{\underline{T}}(t)\underline{P}(t)+\underline{T}(t)\underline{A}(t)\underline{P}(t))' \\ &\quad +\underline{T}(t)\underline{R}(t)\underline{T}'(t)-\hat{\underline{\Sigma}}(t)\underline{P}'(t)\underline{C}_1'(t)\underline{Q}^{-1}(t)\underline{C}_1(t)\underline{P}(t)\hat{\underline{\Sigma}}(t) \end{aligned} \tag{4.5.21}$$

$$\hat{\underline{\Sigma}}(t_o) = \underline{T}(t_o)\underline{\Sigma}_o\underline{T}'(t_o) \quad .$$

The estimate of $\underline{x}(t)$ is given by

$$\hat{\underline{x}}(t) = \underline{P}(t)\hat{\underline{x}}_1(t) + \underline{V}_2(t)\underline{y}_2(t) \quad . \tag{4.5.22}$$

The estimation error covariance matrix is given by

$$\underline{\Sigma}(t) = \underline{P}(t)\hat{\underline{\Sigma}}(t)\underline{P}'(t) \quad . \tag{4.5.23}$$

Therefore

$$\begin{aligned}
\dot{\underline{\Sigma}}(t) &= \dot{\underline{P}}(t)\hat{\underline{\Sigma}}(t)\underline{P}'(t)+\underline{P}(t)\dot{\hat{\underline{\Sigma}}}(t)\underline{P}'(t)+\underline{P}(t)\hat{\underline{\Sigma}}(t)\dot{\underline{P}}'(t) \\
&= (\dot{P}(t)+\underline{P}(t)\dot{\underline{T}}(t)\underline{P}(t)+\underline{P}(t)\underline{T}(t)\underline{A}(t)\underline{P}(t))\hat{\underline{\Sigma}}(t)\underline{P}'(t) \\
&+ \underline{P}(t)\hat{\Sigma}(t)(\dot{\underline{P}}(t)+\underline{P}(t)+\dot{\underline{T}}(t)\underline{P}(t)+\underline{P}(t)\underline{T}(t)\underline{A}(t)\underline{P}(t))' \\
&+ \underline{P}(t)\underline{T}(t)\underline{R}(t)\underline{T}'(t)\underline{P}'(t)-\underline{P}(t)\hat{\underline{\Sigma}}(t)\underline{P}'(t)\underline{C}_1'(t)\underline{Q}^{-1}(t)\underline{C}_1(t)\underline{P}(t)\hat{\underline{\Sigma}}(t)\underline{P}'(t) \\
&= (\underline{A}(t)-\underline{V}_2(t)\tilde{\underline{C}}_2(t))\underline{\Sigma}(t)+\underline{\Sigma}(t)(\underline{A}(t)-\underline{V}_2(t)\tilde{\underline{C}}_2(t))' \\
&+ (\underline{I}_n-\underline{V}_2(t)\underline{C}_2(t))\underline{R}(t)(\underline{I}_n-\underline{V}_2(t)\underline{C}_2(t))'-\underline{\Sigma}(t)\underline{C}_1'(t)\underline{Q}^{-1}(t)\underline{C}_1(t)\underline{\Sigma}(t) \quad .
\end{aligned} \tag{4.5.24}$$

The initial condition is

$$\underline{\Sigma}(t_o) = (\underline{I}_n - \underline{V}_2(t_o)\underline{C}_2(t_o))\underline{\Sigma}_o(\underline{I}_n - \underline{V}_2(t_o)\underline{C}_2(t_o))' \quad . \tag{4.5.25}$$

We note that the error covariance depends on $\underline{V}_2(t)$ which must satisfy

$$\underline{C}_2(t)\underline{V}_2(t) = \underline{I}_{m_2} \quad . \tag{4.5.26}$$

To find the minimum mean square estimates, we have the optimization problem of choosing $\underline{V}_2(t)$ satisfying (4.5.26) and yielding the "least" nonnegative definite $\underline{\Sigma}(t)$. Note that (4.5.24), (4.5.25) is the same as (4.3.25) with

$$\tilde{\underline{L}}_1(t) = \underline{\Sigma}(t)\underline{C}_1'(t)\underline{Q}^{-1}(t) \quad . \tag{4.5.27}$$

One can easily show that the optimum estimator derived by using the Kalman filtering approach is a minimal order optimum observer-estimator. Before comparing the merits of Kalman filtering approach and observer-estimator approach as developed in this chapter, the author would like to point out the falacy of an initiative conception by using the Kalman filtering approach. This is best explained but giving a specific example.

Consider a linear time invariant system described by

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -a_1 & -a_2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ \underline{x}_2(t) \end{bmatrix} + \begin{bmatrix} \xi_1(t) \\ \xi_2(t) \end{bmatrix} \quad . \tag{4.5.28}$$

The observations are

$$y_1(t) = [0 \quad 1]\underline{x}(t) + \eta(t) \tag{4.5.29}$$

$$y_2(t) = [1 \quad 0]\underline{x}(t) \quad . \tag{4.5.30}$$

The noise statistical laws are assumed to be known:

$$E\left\{\int_o^t \underline{\xi}(\tau)d\tau\right\} = \underline{0} \quad ; \quad E\left\{\left(\int_o^t \underline{\xi}(\tau)d\tau\right)\left(\int_o^t \underline{\xi}(\sigma)d\sigma\right)'\right\} = \begin{bmatrix} \tau_1 t & 0 \\ 0 & \tau_2 t \end{bmatrix} \tag{4.5.31}$$

$$E\left\{\int_o^t \eta(\tau)d\tau\right\} = 0 \quad ; \quad E\left\{\left(\int_o^t \eta(\tau)d\tau\right)^2\right\} = qt \quad . \tag{4.5.32}$$

Assume that the estimation process has started at $-\infty$, and our objective now is to find the conditional mean estimate of the state. One "intuitive" argument using the Kalman filtering approach will be as follows. From (4.5.30), we see that we have exact observation in $x_1(t)$, therefore we can assume $x_1(t)$ is known. From (4.5.28) and (4.5.29), we have

$$\dot{x}_2(t) = -a_1x_1(t) - a_2x_2(t) + \xi_2(t) \qquad (4.5.33)$$

$$y_1(t) = x_2(t) + n(t) \quad . \qquad (4.5.34)$$

Since the system is linear and the noises are Gaussian, thus to find the unbiased mean square estimate of $x_2(t)$, we may apply Kalman filter to (4.5.33) and (4.5.34). The error variance, e, in the steady state will satisfy the algebraic equation[50]

$$e^2 + 2a_1qe - r_2q = 0 \quad . \qquad (4.5.35)$$

Therefore the error variance is equal to

$$e = \sqrt{a_1^2q^2 + r_2q} - a_1q > 0 \quad . \qquad (4.5.36)$$

One may make the conclusion that the Kalman filter for (4.5.33) and (4.5.34) will give us the unbiased minimum least square estimates, and the minimum mean square error is given by (4.5.36). Unfortunately, this conclusion is in general <u>false</u>; the reason for this is that the Kalman filter for (4.5.33) and (4.5.34) give us the estimate

$$\tilde{x}_2(t) = E\{x_2(t)|F(y_1(\tau);\tau \in [t_o,t))\} \qquad (4.5.37)$$

whereas the estimate we are looking for is

$$\hat{x}_2(t) = E\{x_2(t)|F(y_2(\tau);\tau \in [t_o,t],y_1(s);s \in [t_o,t))\} \qquad (4.5.38)$$

and in general we have the inclusion of σ-algebra

$$F(y_1(\tau);\tau \in [t_o,t)) \subset F(y_2(\tau);\tau \in [t_o,t],y_1(s);s \in [t_o,t)) \qquad (4.5.39)$$

To proceed with the example, let us define

$$x_{1k}(t) = [-k \quad 1]\underline{x}(t) \tag{4.5.40}$$

where k is an arbitrary number. Using (4.5.30) and (4.5.40), we have

$$\underline{x}(t) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} x_{1k}(t) + \begin{bmatrix} 1 \\ k \end{bmatrix} y_2 \quad . \tag{4.5.41}$$

Taking the derivative of (4.5.40) and using (4.5.28), we have

$$\begin{aligned} \dot{x}_{1k}(t) &= -(k + a_2)x_2(t) - a_1x_1(t) + \xi_2(t) - k\xi_1(t) \\ &= -(k + a_2)x_{1k}(t) - (k(k + a_2) + a_1)y_2(t) + \xi_2(t) - k\xi_1(t) \quad . \end{aligned} \tag{4.5.42}$$

The observation (4.5.29) becomes

$$y_1(t) = x_{1k}(t) + ky_2(t) + \eta(t) \quad . \tag{4.5.43}$$

Define

$$y_{1k}(t) = y_1(t) - ky_2(t) = x_{1k}(t) + \eta(t) \quad . \tag{4.5.44}$$

Since $y_2(s)$, $s \in [t_o, t]$, is known at t, by applying the Kalman filter to (4.5.42) and (4.5.44), we have the steady state error variance, e_k, for the unbiased least square estimate of $x_{1k}(t)$ satisfying the algebraic equation

$$e_k^2 + 2(k + a_2)qe_k - q(k^2r_1 + r_2) = 0 \tag{4.5.45}$$

and so

$$e_k = \sqrt{(k + a_2)^2q^2 + q(k^2r_1 + r_2)} - (k + a_2)q > 0 \quad . \tag{4.5.46}$$

To find the corresponding estimate in $\underline{x}$, we have

$$\underline{\hat{x}}(t) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \hat{x}_{1k}(t) + \begin{bmatrix} 1 \\ k \end{bmatrix} y_2 \qquad (4.5.47)$$

where $\hat{x}_{1k}(t)$ is the estimate given by the Kalman filter for (4.5.42) and (4.5.44). The corresponding error variance for $x_2(t)$ is

$$E\{(x_2(t)-\hat{x}_2(t))^2\} = E\{(x_{1k}(t)-\hat{x}_{1k}(t)+ky_2(t)-ky_2(t))^2\} = E\{(x_{1k}(t)-\hat{x}_{1k}(t))^2\}$$

$$= e_k \quad . \qquad (4.5.48)$$

Clearly $e_k \neq e$ for almost all k ranging from $-\infty$ to ∞. One may then attempt to find the optimum k which give us the smallest e_k. This has easily been carried out by using differential calculus. The optimum value for k^o was found to be:

$$k^o = \frac{a_2^2 q^2}{(q + r_1)^2} + \frac{q r_2}{r_1(q + r_1)} - \frac{a_2 q}{q + r_1} \quad . \qquad (4.5.49)$$

Substituting (4.5.49) into (4.5.46), we have the corresponding error:

$$e^o \triangleq e_{k^o} = \frac{r_1}{q + r_1}\left\{ \sqrt{(qa_2)^2 + \frac{(q + r_1)r_2 q}{r_1}} - qa_2 \right\} \qquad (4.5.50)$$

and clearly we have the strict inequality ($r_1 > 0$, $q > 0$)

$$e^o < \sqrt{(qa_2)^2 + \frac{(q + r_1)}{r_1} r_2 q} - qa_2 < \sqrt{(qa_2)^2 + r_2 q} - qa_2 = e \quad . \qquad (4.5.51)$$

The inequality (4.5.51) indicates that by applying Kalman filter to (4.5.33) and (4.5.34), we <u>do not</u> obtain the best mean square unbiased estimate. We note that the optimum value of k depends on r_1, r_2, and q (assume a_2 is fixed a priori). We may not conclude that the error e^o is <u>the</u> minimum error variance because we only consider a restricted class of transformation in $\underline{x}$ (4.5.40). The only way to check whether e^o is the minimum error variance is to appeal to the projection equation, or equivalently, the Weiner-Hopf equation.

Therefore, we see that conceptually, the Kalman filtering approach is by no means simpler than the observer approach developed in this chapter; because one may find it hard to visualize physically why one transformation of the state is better than the other before the application of a Kalman filter, besides, one may reach false conclusions if one is not careful (see example). Note that one approach is as easy as the other: both involve a deterministic optimization problem, and both need to verify that the derived solution satisfies the projection equation before we can conclude the truly optimum nature of the obtained estimate. In terms of derivation, the Kalman filtering approach is comparatively simpler; but personally, the author thinks that the class of asymptotic unbiased estimator is a more basic conceptual framework to many estimation problems. The observer approach is based precisely on this conception. One distinguishing advantage of using the observer theory approach is that it reveals the detail structural properties of the optimum estimator. This allows us to investigate in detail the asymptotic behavior of the optimum estimator in terms of some intrinsic functional behavior of the system (section 4.4).

(D) Detectability and Observability

We note that observability is a stronger condition than detectability. In section 4.4, we have shown that detectability is a necessary condition for the minimum error covariance, $\underline{\Sigma}^*(t)$, to be uniformly bounded for all $t > t_o$. In the time invariant case, observability is sufficient condition for $\underline{\Sigma}^*(t)$ to be uniformly bounded and for the existence of a steady state value of $\underline{\Sigma}^*(t)$ as $t \to \infty$.

Although a proof is not available yet, it seems very likely that detectability is also sufficient to assure $\underline{\Sigma}^*(t)$ to be uniformly bounded for all $t \geq t_o$.

(E) Estimation in the Presence of Time-Correlated Noise

Let us consider the stochastic system $\mathcal{S}_3^c$ described by

$$\mathcal{S}_3^c: \begin{cases} \dot{\underline{x}}(t) = \underline{A}(t)\underline{x}(t) + \underline{B}(t)\underline{u}(t) + \underline{\xi}(t) & ; \quad \underline{x}(t) \in R^n \\ \underline{y}(t) = \underline{C}(t)\underline{x}(t) + \underline{n}(t) & ; \quad \underline{y}(t) \in R^m \end{cases} \tag{4.5.52}$$

where $\underline{n}(t)$ is a Gaussian Markov process which can be realized by:

$$\dot{\underline{n}}(t) = \tilde{\underline{A}}(t)\underline{n}(t) + \underline{\gamma}(t) \tag{4.5.53}$$

$\underline{x}(t_o)$, $\underline{n}(t_o)$, $\{\underline{\xi}(t),\ t \geq t_o\}$, $\{\underline{\gamma}(t),\ t \geq t_o\}$ are independent statistics. The statistical laws are given by

$$\begin{aligned} \underline{x}(t_o) &\sim \mathcal{G}(\underline{x}_o, \underline{\Sigma}_x^o) \\ \underline{n}(t_o) &\sim \mathcal{G}(\underline{\nu}_o, \underline{\Sigma}_\nu^o) \\ \int_{t_1}^{t_2} \underline{\xi}(t)dt &\sim \mathcal{G}(\underline{0}, \int_{t_1}^{t_2} \underline{R}(t)dt) \\ \int_{t_1}^{t_2} \underline{n}(t)dt &\sim \mathcal{G}(\underline{0}, \int_{t_1}^{t_2} \underline{R}_n(t)dt) \end{aligned} \tag{4.5.54}$$

Define the augmented vectors

$$\underline{x}^a(t) = \begin{bmatrix} \underline{x}(t) \\ \cdots \\ \underline{n}(t) \end{bmatrix} \in R^{n+m} \qquad \underline{\xi}^a(t) = \begin{bmatrix} \underline{\xi}(t) \\ \cdots \\ \underline{\gamma}(t) \end{bmatrix} \in R^{n+m} \tag{4.5.55}$$

and the augmented matrices

$$\underline{A}^a(t) = \begin{bmatrix} \underline{A}(t) & \vdots & \underline{0} \\ \cdots & & \cdots \\ \underline{0} & \vdots & \tilde{\underline{A}}(t) \end{bmatrix} \quad ; \quad \underline{C}^a(t) = [\underline{C}(t) \ \vdots \ \underline{I}_m] \quad ; \quad \underline{B}^a(t) = \begin{bmatrix} \underline{B}(t) \\ \cdots \\ \underline{0} \end{bmatrix} \qquad (4.5.56)$$

We have the equations for the augmented system

$$\mathcal{S}_a^c: \begin{cases} \dot{\underline{x}}^a(t) = \underline{A}^a(t)\underline{x}^a(t) + \underline{B}^a(t)\underline{u}(t) + \underline{\xi}^a(t) \\ \underline{y}(t) = \underline{C}^a(t)\underline{x}^a(t) \end{cases} \qquad (4.5.57)$$

We can apply the derived results to the system $\mathcal{S}_a^c$. Note that the minimal order optimum observer-estimator has dimension n. In the special case when $\underline{C}$ is a constant matrix, we can easily verify that the results obtained agree with those obtained by Bucy [52]. In the general case, the results agree with Bryson and Mehra who considered the problem using the weighted least square approach. Application of the derived results to this special class of problems will be considered in detail in the future.

4.6 Perspective

Qualitative estimation theory for the deterministic system $\mathcal{S}_i^c$ was considered by Luenberger [35], Johnson [49]. Optimum filtering theory for stochastic linear systems was first considered by Wiener [51]. Kalman and Bucy [50] consider the special case of estimating the state of a Gaussian Markov process in the presence of nondegenerate Gaussian white observation noise. Estimation in the singular situation (i.e., noise free observation) was considered by Root [54]. (See also Van Trees [47] for detailed bibliography.) Estimation in the presence of colored noise only was considered by Bucy [52], Mehra and Bryson [53], Geesey and Kailath [55]. The consideration in this chapter provides a unifying approach to linear estimation problems in general. This approach is valuable in the way that it reveals the intrinsic structural properties of the estimating device.

The asymptotic behavior of the estimator was not investigated in detail in the literature save for the case of estimation in the presence of nondegenerative Gaussian white observation noise [50]. In this special case, the asymptotic behavior of the optimum estimator was investigated through its dual relation with an optimal regulation problem [50]. The investigation in section 4.4 is original. In this contribution, we can study, in all general situations, the asymptotic behavior of optimum estimators; it also provides the concepts required for qualitative estimation theory for deterministic linear systems.

CHAPTER V

OPTIMAL CONTROL OF STOCHASTIC LINEAR SYSTEMS WITH KNOWN DYNAMICS

5.1 Introduction

In this chapter, we are mainly concerned with the problem of controlling a linear system with known dynamics, under the assumption that perfect information is not available. To have the problem be completely general, we assume that there are unknown driving disturbances, and partial observation of the state in the presence (or absence) of observation noise. A special case of the problem was investigated by Joseph and Tou [56], Gunckel and Franklin [58], Wonham [22], [27], where they assumed that the observation noise is a nondegenerate white Gaussian process. In our investigation, we assume that the observation noise is in general degenerate without loss of generality; we shall model the problem as one where some output variables can be observed perfectly (noise-free) while the others are observed in the presence of white Gaussian noise. This general information includes [56], [58], [27] as special cases.

The structure of this chapter is as follows. In section 5.2, we consider the estimation problem for the discrete case where we are allowed to use feedback control. Using the results in chapter 3, we derive a stochastic difference equation for the conditional mean estimates of the current state. In section 5.3, we shall state the stochastic control problem and the optimality criteria is used to verify the optimal solution. The general results are then applied to a special case where the observation noise is sequentially correlated. In section 5.4 and section 5.5 we treat the continuous analog of section 5.2 and 5.3. The results

can be summarized as the Separation Theorem. In section 5.6, discussions of results and indication of some further research is given.

5.2 Estimation with Feedback for Discrete Linear Systems

Consider a discrete linear system $\mathcal{S}_3$ described by

$$\mathcal{S}_3: \quad \begin{aligned} \underline{x}(k+1) &= \underline{A}(k)\underline{x}(k) + \underline{B}(k)\underline{u}(k) + \underline{\xi}(k) \\ \underline{y}(k) &= \underline{C}(k)\underline{x}(k) + \underline{\eta}(k) \end{aligned} \tag{5.2.1}$$

where $\underline{x}(k) \in R^n$, $\underline{u}(k) \in R^r$, $\underline{y}(k) \in R^m$. $\underline{x}(0)$, $\underline{\xi}(k)$, $\underline{\eta}(k)$, $k = 0, 1, \ldots$ are independent Gaussian random vectors with statistical law given by (3.3.2) to (3.3.4). The control $\underline{u}(k)$ is feedback in nature. Let us denote the control sequence by

$$U(i,j) \triangleq \{\underline{u}(i), \underline{u}(i+1), \ldots, \underline{u}(j)\} \qquad i > j \quad . \tag{5.2.2}$$

The observation statistic at time k is $\underline{y}_{U(o,k-1)}(k)$, where the subscript $U(o,k-1)$ is to indicate that the past control sequence, $U(0,k-1)$, has been applied to the system. The accumulative observation statistic at time k is

$$Y_{U(o,k-1)}(k) = \{\underline{y}(0), \underline{y}_{\underline{u}(o)}(1), \ldots, \underline{y}_{U(o,k-1)}(k)\} \quad . \tag{5.2.3}$$

We assume that the control is of the form

$$\underline{u}(k) = \underline{\phi}(k, Y_{U(o,k-1)}(k)) \qquad k = 0, 1, \ldots \tag{5.2.4}$$

where $\underline{\phi}(k,\cdot)$ is a measurable function from $F(Y_{U(o,k-1)})$ to R^r. (5.2.4) implies that the control is a function of past accumulative observation information. In the following, we shall denote $F(Y_{U(o,k-1)}(k))$ by $F(k,U(0,k-1))$.

The information revealed by the accumulative information at time k about the dynamical state of the system is contained in the sub-σ-algebra $F(k,U(0,k-1))$. For some control purposes, the detailed knowledge of $F(k,U(0,k-1))$ is sufficient but not necessary. In most cases, since the knowledge <u>about</u> the present state is necessary and sufficient for designing a feedback control strategy, then the knowledge of the conditional distribution of the present state, $\underline{x}(k)$, is necessary and sufficient (see chapter 2, section 2.2). In the following, we shall prove that the conditional distribution of $\underline{x}(k)$ can be parameterized by some finite dimensional quantities.

<u>Theorem 5.2.1</u>: For the system S_3 where $\underline{u}(k)$ is of the form (5.2.4), the conditional distribution of $\underline{x}(k)$ is a Gaussian random vector, and so is parameterized by its conditional mean, $\hat{\underline{x}}(k|k)$, and conditional covariance $\underline{\Sigma}(k)$ which are given by $(k = 0, 1, \ldots)$

$$\mathcal{E}_T^3: \begin{cases} \underline{z}(k+1) = \underline{T}(k+1)\underline{A}(k)\underline{P}(k)\underline{z}(k)+\underline{T}(k+1)\underline{A}(k)\underline{V}^*(k)\underline{y}(k)+\underline{T}(k+1)\underline{B}(k)\underline{u}(k) \\ \hat{\underline{x}}(k|k) = \underline{P}(k)\underline{z}(k)+\underline{V}^*(k)\underline{y}(k) \end{cases} \tag{5.2.5}$$

$$\begin{aligned} \underline{\Sigma}(k+1) &= \underline{\Delta}(k) - \underline{V}^*(k+1)\underline{C}(k+1)\underline{\Delta}(k) \\ \underline{\Sigma}(0) &= \underline{\Sigma}_o - \underline{V}^*(0)\underline{C}(0)\underline{\Sigma}_o \end{aligned} \tag{5.2.6}$$

where

$$\underline{\Delta}(k) \triangleq \underline{A}(k)\underline{\Sigma}(k)\underline{A}'(k)+\underline{R}(k) \quad ; \quad \underline{V}^*(0) = \underline{\Sigma}_o\underline{C}'(0)(\underline{C}(0)\underline{\Sigma}_o\underline{C}'(0)+\underline{Q}(0))^{-1} \tag{5.2.7}$$

and

$$\underline{V}^*(k+1) \in \mathcal{V}_k(\underline{\Sigma}(k)) = \{\underline{V} \in M_{nm} | \underline{V}[\underline{C}(k+1)\underline{\Delta}(k)\underline{C}'(k+1)+\underline{Q}(k)] = \underline{\Delta}(k)\underline{C}'(k+1)\}$$

$\underline{P}(t)$, $\underline{T}(t)$ satisfy the relation $\underline{P}(k)\underline{T}(k) + \underline{V}^*(k)\underline{C}(k) = \underline{I}_n$.

Proof: Let us break $\underline{x}(k)$ into two vectors:

$$\underline{x}(k) = \underline{x}_1(k) + \underline{x}_2(k) \tag{5.2.8}$$

where $\underline{x}_1(k)$, $\underline{x}_2(k) \in R^n$ are given by

$$\underline{x}_2(k+1) = \underline{A}(k)\underline{x}_2(k) + \underline{\xi}(k) \quad ; \quad \underline{x}_2(0) = \underline{x}(0) \tag{5.2.9}$$

$$\underline{x}_1(k+1) = \underline{A}(k)\underline{x}_1(k) + \underline{B}(k)\underline{u}(k) \quad ; \quad \underline{x}_1(0) = \underline{0} \tag{5.2.10}$$

and $\underline{u}(k)$ is of the form (5.2.4). From (5.2.10), we note that $\{\underline{x}_1(i)\}_{i=0}^{k}$ is $F(k,U(0,k-1))$-measurable, and so we have from (5.2.8) that

$$\hat{\underline{x}}(k|k) = E\{\underline{x}_2(k)|F(k,U(0,k-1))\} + \underline{x}_1(k) \quad . \tag{5.2.11}$$

Let us define

$$\underline{y}_2(k) = \underline{y}(k) - \underline{C}(k)\underline{x}_1(k) = \underline{C}(k)\underline{x}_2(k) + \underline{n}(k) \tag{5.2.12}$$

and define $F_2(k) \triangleq F(\underline{y}_2(0),\ldots,\underline{y}_2(k))$. $\{\underline{y}(i)\}_{i=0}^{k}$ and $\{\underline{x}_1(i)\}_{i=0}^{k}$ are $F(U(0,k-1),k)$-measurable, so $\{\underline{y}_2(i)\}_{i=0}^{k}$ is $F(k,U(0,k-1))$-measurable; therefore

$$F_2(k) \subset F(k,U(0,k-1)) \quad . \tag{5.2.13}$$

Using (5.2.4), (5.2.10) and (5.2.12) we have

$$\underline{x}_1(k+1) = \underline{A}(k)\underline{x}_1(k)+\underline{B}(k)\underline{\phi}(k,\underline{y}_1(0)+\underline{C}(0)\underline{x}_1(0),\ldots,\underline{y}_1(k)+\underline{C}(k)\underline{x}_1(k)) \quad ;$$
$$\underline{x}_1(0) = \underline{0} \quad . \tag{5.2.14}$$

Inductively, we have $\{\underline{x}_1(i)\}_{i=0}^{k}$ is $F_2(k)$-measurable, and so from (5.2.12), $\underline{y}(k)$ is $F_2(k)$-measurable. We have then

$$F(k,\underline{U}(0,k-1)) \subset F_2(k) \tag{5.2.15}$$

(5.2.13) and (5.2.14) imply that

$$F_2(k) = F(k,\underline{U}(0,k-1)) \quad . \tag{5.2.16}$$

Let us define

$$\hat{\underline{x}}_2(k|k) = E\{\underline{x}_2(k)|F_2(k)\} \tag{5.2.17}$$

(5.2.11) and (5.2.15) give

$$\hat{\underline{x}}(k|k) = \hat{\underline{x}}_2(k|k) + \underline{x}_1(k) \quad . \tag{5.2.18}$$

Now consider the stochastic system $\mathcal{S}_2$ and the deterministic system $\mathcal{S}_1$ described by

$$\mathcal{S}_2: \begin{cases} \underline{x}_2(k+1) = \underline{A}(k)\underline{x}_2(k) + \underline{\xi}(k) \; ; \quad \underline{x}_2(0) = \underline{x}(0) \sim G(\underline{x}_o,\underline{\Sigma}_o) \\ \underline{y}_2(k) = \underline{C}(k)\underline{x}_2(k) + \underline{n}(k) \end{cases} \tag{5.2.19}$$

$$\mathcal{S}_1: \begin{cases} \underline{x}_1(k+1) = \underline{A}(k)\underline{x}(k) + \underline{B}(k)\underline{u}(k) \; ; \quad \underline{x}_1(0) = \underline{0} \\ \underline{y}_1(k) = \underline{C}(k)\underline{x}_1(k) \end{cases} \tag{5.2.20}$$

Since $\underline{x}_1(0)$ is known exactly, $\underline{x}_1(k)$ can be reconstructed by <u>any</u> class of observers. The conditional distribution of $\underline{x}_2(k)$ given $F_2(k)$ is Gaussian, the conditional mean, $\hat{\underline{x}}_2(k|k)$, and the conditional covariance, $\underline{\Sigma}(k)$, are given by: (Chapter 3, section 3.2 and section 3.3)

$$\mathcal{E}_T^2: \begin{cases} \underline{z}_2(k+1) = \underline{T}(k+1)\underline{A}(k)\underline{P}(k)\underline{z}_2(k)+\underline{T}(k+1)\underline{A}(k)\underline{V}^*(k)\underline{y}_2(k) \; ; \quad \underline{z}_1(0) = \underline{T}(0)\underline{x}_o \\ \hat{\underline{x}}_2(k|k) = \underline{P}(k)\underline{z}_2(k)+\underline{V}^*(k)\underline{y}_2(k) \end{cases} \tag{5.2.21}$$

and

$$\underline{\Sigma}(k+1) = \underline{\Delta}(k) - \underline{V}^*(k+1)\underline{C}(k+1)\underline{\Delta}(k)$$

$$\underline{\Sigma}(0) = \underline{\Sigma}_o - \underline{\Sigma}_o\underline{C}'(0)[\underline{C}(0)\underline{\Sigma}_o\underline{C}'(0) + \underline{Q}(0)]^{-1}\underline{C}(0)\underline{\Sigma}_o \tag{5.2.22}$$

$$\underline{\Delta}(k) \triangleq \underline{A}(k)\underline{\Sigma}(k)\underline{A}'(k) + \underline{R}(k) \tag{5.2.23}$$

$$\underline{V}^*(0) = \underline{\Sigma}_o\underline{C}'(0)(\underline{C}(0)\underline{\Sigma}_o\underline{C}'(0)+\underline{Q}(0))^{-1} \quad ; \quad \underline{V}^*(k+1) \in \mathcal{V}_k(\underline{\Sigma}(k)) \, ; \, k = 0, 1, \ldots \tag{5.2.24}$$

Construct $\underline{x}_1(k)$ by using

$$\mathcal{E}_T^1: \begin{array}{l} \underline{z}_1(k+1) = \underline{T}(k+1)\underline{A}(k)\underline{P}(k)\underline{z}_1(k)+\underline{T}(k+1)\underline{A}(k+1)\underline{V}^*(k)\underline{y}_1(k)+\underline{T}(k+1)\underline{B}(k)\underline{u}(k) \\ \underline{x}_1(k) = \underline{P}(k)\underline{z}_1(k)+\underline{V}^*(k)\underline{y}_2(k) \quad ; \quad \underline{z}_1(0) = \underline{0} \end{array} \tag{5.2.25}$$

where $\{\underline{V}^*(k)\}_{k=0}^{\infty}$ is given by (5.2.21) to (5.2.24). From (5.2.19), (5.2.20) and (5.2.8), we have

$$\underline{y}(k) = \underline{y}_1(k) + \underline{y}_2(k) \quad . \tag{5.2.26}$$

Define the vector

$$\underline{z}(k) = \underline{z}_1(k) + \underline{z}_2(k) \quad . \tag{5.2.27}$$

By equations (5.2.18), (5.2.21) to (5.2.26), we have the conditional mean estimate of $\underline{x}(k)$ generated by

$$\mathcal{E}_T^3: \begin{array}{l} \underline{z}(k+1) = \underline{T}(k+1)\underline{A}(k+1)\underline{P}(k)\underline{z}(k)+\underline{T}(k+1)\underline{A}(k)\underline{V}^*(k)\underline{y}(k)+\underline{T}(k+1)\underline{B}(k)\underline{u}(k) \\ \hat{\underline{x}}(k|k) = \underline{P}(k)\underline{z}(k)+\underline{V}^*(k)\underline{y}(k) \quad ; \quad \underline{z}(0) = \underline{T}(0)\underline{x}_o \end{array} \tag{5.2.28}$$

with $\{\underline{V}^*(k)\}_{k=0}^{\infty}$ given by (5.2.21) to (5.2.24).

Using equations (5.2.8) and (5.2.18), we have

$$\underline{\hat{x}}(k|k) - \underline{x}(k) = \underline{\hat{x}}_2(k|k) + \underline{x}_1(k) - \underline{x}_1(k) - \underline{x}_2(k) = \underline{\hat{x}}_2(k|k) - \underline{x}_2(k) \tag{5.2.29}$$

$\{\underline{\Sigma}(k)\}_{k=0}^{\infty}$ given by (5.2.21) to (5.2.24) is the conditional covariance of $\underline{x}_2(k)$, and so it is also the conditional covariance of $\underline{x}(k)$. Since $\underline{x}_1(k)$ is $F(k,U(0,k-1))$-measurable, (5.2.8) implies that the conditional distribution of $\underline{x}(k)$ is Gaussian, by virtue that the conditional distribution of $\underline{x}_1(k)$ is Gaussian.

We note from (5.2.3) and (5.2.4) that the accumulative statistic at time k depends on the control chosen which in turn depends on past accumulative statistics. But as long as we are interested in the present state of the system, the information contained in $F(k,U(0,k-1))$ about $\underline{x}(k)$ is equivalent in some sense to the statistical information contained in the conditional distribution of $\underline{x}(k)$. Theorem 5.2.1 says that the conditional distribution of $\underline{x}(k)$ is Gaussian, and thus all the statistical information revealed by accumulated observation statistics is summarized by the conditional mean, $\underline{\hat{x}}(k|k)$, and conditional covariance, $\underline{\Sigma}(k)$. From (5.2.21) to (5.2.23), we see that $\underline{\Sigma}(k)$ can be precomputed before any observation is made and any control is applied. Therefore, all the statistical information about the state at time k is summarized in the random vector $\underline{\hat{x}}(k|k)$.

5.3 Stochastic Control of Discrete Linear Systems with Quadratic Criteria

In this section, we shall consider the problem of controlling the discrete linear system $\mathcal{S}_3$ with quadratic criteria:

$$J(\underline{u}) = E\left\{\underline{x}'(N)\underline{F}\,\underline{x}(N) + \sum_{k=0}^{N-1}\left(\underline{x}'(k)\underline{W}(k)\underline{x}(k) + \underline{u}'(k)\underline{M}(k)\underline{u}(k)\right)\right\} \tag{5.3.1}$$

with $\underline{F} \geq \underline{0}$, $\underline{W}(k) \geq \underline{0}$, and $\underline{M}(k) > \underline{0}$. We are to find a control law of the form (5.2.4) which will minimize (5.3.1) subject to (5.2.1).

Using lemma 2.2.6 and (5.2.4), the cost $J(\underline{u})$ can be rewritten as:

$$J(\underline{u}) = E\Big\{E\{\underline{x}'(N)\underline{F}\ \underline{x}(N)|F(N,U(0,N-1)\} + \sum_{k=0}^{N-1} E\{\underline{x}'(k)\underline{W}(k)\underline{x}(k)+\underline{u}'(k)\underline{M}(k)\underline{u}(k).$$

$$F(k,U(0,k-1))\}\Big\}$$

$$= E\Big\{\hat{\underline{x}}'(N|N)\underline{F}\ \hat{\underline{x}}(N|N) + \sum_{k=0}^{N-1}\Big(\hat{\underline{x}}'(k|k)\underline{W}(k)\hat{\underline{x}}(k|k)+\underline{u}'(k)\underline{M}(k)\underline{u}(k)\Big)\Big\}$$

$$+\mathrm{tr}\Big(\underline{F}\ \underline{\Sigma}(N) + \sum_{k=0}^{N-1} \underline{W}(k)\underline{\Sigma}(k)\Big) \qquad (5.3.2)$$

where $\{\underline{\Sigma}(j)\}_{j=0}^{N}$ is given by (5.2.22) to (5.2.24). Since $\{\underline{\Sigma}(j)\}_{j=0}^{N}$ is independent of the control, minimizing (5.3.1) is equivalent to minimizing

$$J'(\underline{u}) = E\Big\{\hat{\underline{x}}'(N|N)\underline{F}\ \hat{\underline{x}}(N|N) + \sum_{k=0}^{N-1}\Big(\hat{\underline{x}}'(k|k)\underline{W}(k)\hat{\underline{x}}(k|k) + \underline{u}'(k)\underline{M}(k)\underline{u}(k)\Big)\Big\} \quad . \qquad (5.3.3)$$

From (5.2.5), the equation for $\hat{\underline{x}}(k|k)$ is given by

$$\hat{\underline{x}}(k+1|k+1) = \underline{A}(k)\hat{\underline{x}}(k|k) - \underline{V}^*(k+1)\underline{C}(k+1)\underline{A}(k)(\hat{\underline{x}}(k|k) - \underline{x}(k)) + \underline{B}(k)\underline{u}(k)$$

$$+ \underline{V}^*(k+1)\underline{C}(k+1)\underline{\xi}(k) + \underline{V}^*(k+1)\underline{n}(k+1) \qquad (5.3.4)$$

where $\underline{V}^*(k)$, $k = 0, 1, \ldots, N$ are given by (5.2.21) to (5.2.23). The process $\{\underline{x}(k)\}_{k=0}^{N}$ is given by (5.2.1). We have now a stochastic problem to solve: <u>Find a control law of the form (5.2.4) such that the cost (5.3.3) is minimized subject to the constraints (5.3.4) and (5.3.1)</u>.

<u>Lemma 5.3.1</u>: The control law

$$\underline{u}^*(k) = -(\underline{M}(k)+\underline{B}'(k)\underline{K}(k+1)\underline{B}(k))^{-1}\underline{B}'(k)\underline{K}(k+1)\underline{A}(k)\hat{\underline{x}}(k|k) \qquad (5.3.5)$$

$$\underline{K}(k) = \underline{A}(k)(\underline{K}(k+1)-\underline{K}(k+1)\underline{B}(k)(\underline{M}(k)+\underline{B}'(k)\underline{K}(k+1)\underline{B}(k))^{-1}\underline{B}'(k)\underline{K}(k+1))$$

$$\underline{A}(k)+\underline{W}(k) \quad ; \quad \underline{K}(N) = \underline{F} \qquad (5.3.6)$$

is the optimal control law to the above stochastic control problem, i.e., let $\{\underline{u}(k)\}_{k=0}^{N}$ be any control law of the form (5.2.4), we have

$$J'(\underline{u}^*) \le J'(\underline{u}) \quad . \tag{5.3.7}$$

The optimal cost-to-go is given by

$$J_*'(k,\underline{\hat{x}}) \triangleq E\left\{\underline{\hat{x}}^{*\prime}(N|N)E\,\underline{\hat{x}}^*(N|N)+\sum_{i=k}^{N-1}\underline{\hat{x}}^{*\prime}(i|i)W(i)\underline{\hat{x}}^*(i|i)+\underline{u}^{*\prime}(i)\underline{M}(i)\underline{u}^*(i)\,|\,\underline{\hat{x}}^*(k|k)=\underline{\hat{x}}\right\}$$

$$= \underline{\hat{x}}'\underline{K}(k)\underline{\hat{x}}+\mathrm{tr}\sum_{i=k}^{N-1}(\underline{\Delta}(i)-\underline{\Sigma}(i+1))\underline{K}(i+1) \quad . \tag{5.3.8}$$

<u>Proof:</u> We shall prove the lemma by using the Optimality Criterion (theorem 2.4.3). Let us define for $k = 0, 1, \ldots, N$

$$C(k,\underline{\hat{x}}) \triangleq \underline{\hat{x}}'\underline{K}(k)\underline{\hat{x}} + \mathrm{tr}\sum_{i=k}^{N-1}(\underline{\Delta}(i) - \underline{\Sigma}(i+1))\underline{K}(i+1)$$

$$= \underline{\hat{x}}'K(k)\underline{\hat{x}} + \mathrm{tr}\sum_{i=k}^{N-1}(\underline{V}^*(i+1)\underline{C}(i+1)\underline{\Delta}(i)\underline{K}(i+1) \tag{5.3.9}$$

where $\{\underline{K}(k)\}_{k=0}^{N}$ satisfies (5.3.6). We have from (5.3.6)

$$C(N,\underline{\hat{x}}) = \underline{\hat{x}}'\underline{F}\,\underline{\hat{x}} \quad . \tag{5.3.10}$$

Let $U(0,k-1)$ be arbitrary control sequence, and denote

$$\underline{\hat{x}} = E\{\underline{\hat{x}}(k|k)|F(k,U(0,k-1)\} = E\{\underline{x}(k)|F(k,U(0,k-1))\} \quad . \tag{5.3.11}$$

Let

$$\underline{\hat{x}}^*(k+1) = \underline{A}(k)\underline{\hat{x}} - \underline{V}(k+1)\underline{C}(k+1)\underline{A}(k)(\underline{\hat{x}}-\underline{x}(k)) + \underline{V}^*(k+1)\underline{C}(k+1)\underline{\xi}(k)$$
$$+ \underline{V}(k+1)\underline{n}(k) + \underline{B}(k)\underline{u}^*(k) \tag{5.3.12}$$

$$\underline{\hat{x}}^o(k+1) = \underline{A}(k)\underline{\hat{x}} - \underline{V}(k+1)\underline{C}(k+1)\underline{A}(k)(\underline{\hat{x}}-\underline{x}(k)) + \underline{V}^*(k+1)\underline{C}(k+1)\underline{\xi}(k)$$
$$+ \underline{V}(k+1)\underline{n}(k) + \underline{B}(k)\underline{u}^o(k) \tag{5.3.13}$$

where $\underline{u}^*(k)$ is given by (5.3.5) with $\hat{\underline{x}}$ replacing $\hat{\underline{x}}(k|k)$, and $\underline{u}^o(k)$ is $F(k,U(0,k-1))$-measurable function. We have from (5.3.12), (5.3.5), (5.3.6):

$$E\{\hat{\underline{x}}'\underline{W}(k)\hat{\underline{x}}+\underline{u}^{*\prime}(k)\underline{M}(k)\underline{u}^*(k)|F(k,U(0,k-1)\} = \hat{\underline{x}}'\underline{K}(k)\hat{\underline{x}}-\hat{\underline{x}}'\underline{A}'(k)\underline{K}(k+1)\underline{A}(k)\hat{\underline{x}}$$
$$-\underline{u}^{*\prime}(k)\underline{B}'(k)\underline{K}(k+1)\underline{A}(k)\hat{\underline{x}}+\underline{u}^{*\prime}(k)\underline{M}(k)\underline{u}^*(k)$$
$$= \hat{\underline{x}}'\underline{K}(k)\hat{\underline{x}}(k)-\hat{\underline{x}}^{*\prime}(k+1)\underline{K}(k+1)\hat{\underline{x}}^*(k+1)+\underline{u}^{*\prime}(k)(\underline{M}(k)+\underline{B}'(k)\underline{K}(k)\underline{B}(k))\underline{u}^*(k)$$
$$+\hat{\underline{x}}'\underline{A}'(k)\underline{K}(k+1)\underline{B}(k)\underline{u}^*(k)+tr\{\underline{V}(k+1)\underline{C}(k+1)\underline{\Delta}(k)\underline{K}(k+1)\}$$
$$= \hat{\underline{x}}'\underline{K}(k)\hat{\underline{x}}(k)-\hat{\underline{x}}^{*\prime}(k+1)\underline{K}(k+1)\hat{\underline{x}}^*(k)+tr\{(\underline{\Delta}(k)-\underline{\Sigma}(k+1))\underline{K}(k+1)\} \quad . \tag{5.3.14}$$

Combining (5.3.9) and (5.3.14) we have

$$E\{\hat{\underline{x}}'\underline{W}(k)\hat{\underline{x}}(k)+\underline{u}^{*\prime}(k)\underline{M}(k)\underline{u}^*(k)+C(k+1,\hat{\underline{x}}^*(k+1)|F(k,U(0,k-1))\}-C(k,\hat{\underline{x}}) = \underline{0} \quad . \tag{5.3.15}$$

Since $\underline{u}^o(k)$ is $F(U(0,k-1),k)$-measurable, we have from (5.3.13), (5.3.5), (5.3.6):

$$E\{\hat{\underline{x}}'\underline{W}(k)\hat{\underline{x}}+\underline{u}^{o\prime}(k)\underline{M}(k)\underline{u}^o(k)|F(k,U(0,k-1))\} = \hat{\underline{x}}'\underline{K}(k)\hat{\underline{x}}-\hat{\underline{x}}'\underline{A}'(k)\underline{K}(k+1)\underline{A}(k)\hat{\underline{x}}$$
$$+\hat{\underline{x}}'\underline{A}'(k)\underline{K}(k+1)\underline{B}(k)(\underline{M}(k)+\underline{B}'(k)\underline{K}(k+1)\underline{B}(k))^{-1}\underline{B}'(k)\underline{K}(k+1)\underline{A}(k)\hat{\underline{x}}+\underline{u}^{o\prime}(k)\underline{M}(k)\underline{u}^o(k)$$
$$= \hat{\underline{x}}'\underline{K}(k)\hat{\underline{x}}-\hat{\underline{x}}^{*\prime}(k+1)\underline{K}(k+1)\hat{\underline{x}}^o(k+1)+\underline{u}^{o\prime}(k)(\underline{M}(k)+\underline{B}'(k)\underline{K}(k+1)\underline{B}(k))\underline{u}^o(k)$$
$$+\underline{u}^{o\prime}(k)(\underline{M}(k)+\underline{B}'(k)\underline{K}(k+1)\underline{B}(k))\underline{u}^*(k)+\underline{u}^{*\prime}(k)(\underline{M}(k)+\underline{B}'(k)\underline{K}(k+1)\underline{B}(k))\underline{u}^o(k)$$
$$+\underline{u}^*(k)(\underline{M}(k)+\underline{B}'(k)\underline{K}(k+1)\underline{B}(k))\underline{u}^*(k)+tr\{(\underline{\Delta}(k)-\underline{\Sigma}(k+1)\underline{K}(k+1))\}$$
$$= \hat{\underline{x}}'\underline{K}(k)\hat{\underline{x}}-\hat{\underline{x}}^o(k+1)\underline{K}(k+1)\hat{\underline{x}}^o(k+1)+(\underline{u}^o(k)-\underline{u}^*(k))'(\underline{M}(k)+\underline{B}'(k)\underline{K}(k+1)\underline{B}(k))(\underline{u}^o(k)$$
$$-\underline{u}^*(k))+tr\{\underline{\Delta}(k)-\underline{\Sigma}(k+1)\underline{K}(k+1)\} \quad . \tag{5.3.16}$$

Combining (5.3.9), (5.3.16) and (5.3.15) we have since $\underline{M}(k) > \underline{0}$ and $\underline{K}(k+1) \geq \underline{0}$:

$$0 = E\{\underline{\hat{x}}'\underline{W}(k)\underline{\hat{x}}+\underline{u}^{*\prime}(k)\underline{M}(k)\underline{u}^{*}(k)+C(k+1,\underline{\hat{x}}^{*}(k+1)|F(k,U(0,k-1))\}-C(k,\underline{\hat{x}})$$
$$\leq E\{\underline{\hat{x}}'\underline{W}(k)\underline{\hat{x}}+\underline{u}^{o\prime}(k)\underline{M}(k)\underline{u}^{o}(k)+C(k+1,\underline{\hat{x}}^{o}(k+1)|F(k,U(0,k-1))\}-C(k,\underline{\hat{x}}) \quad . \tag{5.3.17}$$

The lemma follows from the Optimality Criterion.

Theorem 5.3.2: The control law $\underline{u}^*(k)$, $k = 0, 1, \ldots, N$ given by (5.3.5) and (5.3.6) is the optimal control law which minimizes the cost (5.3.1) subject to (5.2.1) and (5.2.4). The optimal cost to go can be expressed as:

$$J_*(k,\underline{\hat{x}}) = \underline{\hat{x}}'\underline{K}(k)\underline{\hat{x}} + \mathrm{tr}\sum_{i=k}^{N-1}[(\underline{\Delta}(i) - \underline{\Sigma}(i+1))\underline{K}(i+1) + \underline{W}(i)\underline{\Sigma}(i)] + \underline{F}\ \underline{\Sigma}(N) \quad . \tag{5.3.18}$$

This follows trivially from lemma 5.3.1 and equation (5.3.2).

Note that $\underline{\Sigma}(k)$, $\underline{\Delta}(k)$, $\underline{K}(k)$, $k = 0, 1, \ldots, N$ can all be precomputed when the noises distribution laws and the weightings ($\underline{F}$, $\underline{W}(k)$, $\underline{M}(k)$) are all given. The performance measure can be easily evaluated when the conditional mean of the state vector is computed via a minimal order optimum observer-estimator. From (5.2.27) and (5.3.5), we see that the optimal control law can be written as

$$\underline{u}^*(k) = -(\underline{M}(k) + \underline{B}'(k)\underline{K}(k+1)\underline{B}(k))^{-1}\underline{B}'(k)\underline{K}(k+1)\underline{A}(k)\underline{P}(k)\underline{z}(k)$$
$$-(\underline{M}(k) + \underline{B}'(k)\underline{K}(k+1)\underline{B}(k))^{-1}\underline{B}'(k)\underline{K}(k+1)\underline{A}(k)\underline{V}^*(k)\underline{y}(k) \quad . \tag{5.3.19}$$

Denote the pure feedback portion of $\underline{u}^*(k)$ by

$$\underline{u}_1^*(k) = -(\underline{M}(k) + \underline{B}'(k)\underline{K}(k+1)\underline{B}(k))^{-1}\underline{B}'(k)\underline{K}(k+1)\underline{A}(k)\underline{V}^*(k)\underline{y}(k) \tag{5.3.20}$$

and feedback after compensation

$$\underline{u}_2^*(k) = -(\underline{M}(k) + \underline{B}'(k)\underline{K}(k+1)\underline{B}(k))^{-1}\underline{B}'(k)\underline{K}(k+1)\underline{A}(k)\underline{P}(k)\underline{z}(k) \quad . \tag{5.3.21}$$

The optimal control is composed of:

$$\underline{u}^*(k) = \underline{u}_1^*(k) + \underline{u}_2^*(k) \quad . \tag{5.3.19'}$$

The detail structure of the optimal control system is described in Figure 5.1.

When the observation noise is nondegenerate, i.e., $\underline{Q}(k) > \underline{0}$, $k = 0, 1, \ldots$, we have the usual separation results first derived by Joseph and Tou. Theorem 5.3.2 indicates that separation is true under more general assumptions when $\underline{Q}(k)$ and $\underline{R}(k)$ are nonnegative definite and even when they are both zero matrices. The theorem can also be applied to the case when the observation noise is sequentially correlated. In the following, we shall treat this special case in some detail.

Consider the system $\tilde{\mathcal{S}}_2$ described by

$$\tilde{\mathcal{S}}_2: \begin{array}{l} \underline{x}(k+1) = \underline{A}(k)\underline{x}(k) + \underline{B}(k)\underline{u}(k) + \underline{\xi}(k) \\ \underline{y}(k) = \underline{C}(k)\underline{x}(k) + \underline{n}(k) \end{array} \tag{5.3.22}$$

$\{\underline{n}(k)\}_{k=0}^{\infty}$ is sequentially correlated and is described by

$$\underline{n}(k+1) = \underline{\tilde{A}}(k)\underline{n}(k) + \underline{\gamma}(k) \quad . \tag{5.3.23}$$

We shall assume that $\underline{\xi}(k)$, $\underline{\gamma}(k)$, $k = 0, 1, \ldots$, $\underline{x}(0)$ and $\underline{n}(0)$ are independent Gaussian random vectors with statistical laws given by (3.3.2), (3.3.3) and (3.7.14). The control problem is to find control $\underline{u}^*(k)$ of the form (5.2.4) which will minimize the cost (5.3.1) subject to (5.3.22), (5.3.23). From (5.3.22) and (5.3.23) we have the augmented system

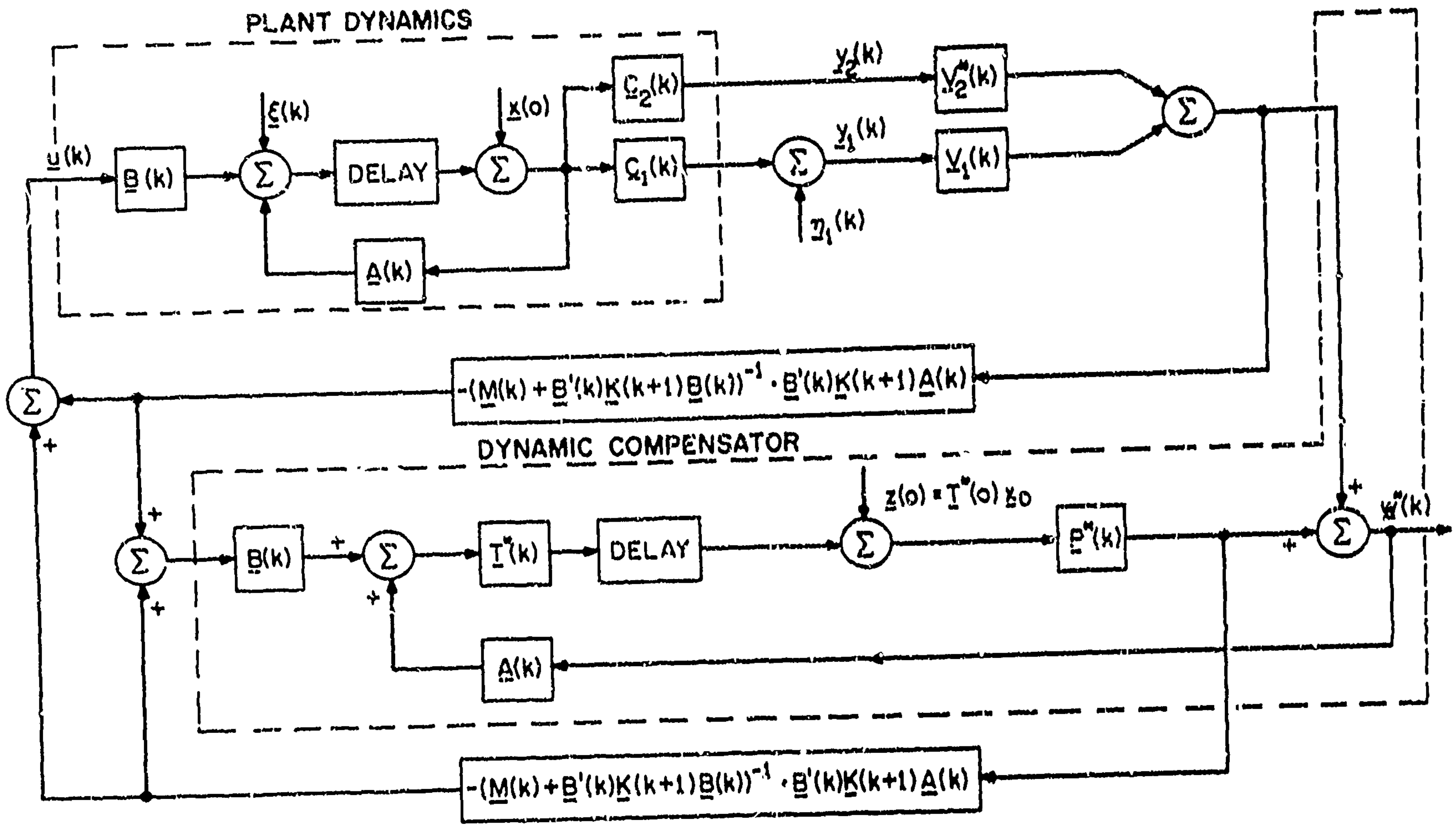


Fig. 5.1 DISCRETE-TIME STOCHASTIC LINEAR OPTIMAL CONTROL SYSTEM. SOME OBSERVATION CHANNELS ARE NOISE-FREE.

$$S_a: \quad \begin{aligned} \underline{x}^a(k+1) &= \underline{A}^a(k)\underline{x}^a(k) + \underline{B}^a(k)\underline{u}(k) + \underline{\xi}^a(k) \\ \underline{y}^a(k) &= \underline{C}^a(k)\underline{x}^a(k) \end{aligned} \tag{5.3.24}$$

where

$$\underline{x}^a(k) = \begin{bmatrix} \underline{x}(k) \\ \cdots \\ \underline{r}(k) \end{bmatrix} \; ; \quad \underline{A}^a(k) = \begin{bmatrix} \underline{A}(k) & \vdots & \underline{0} \\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{A}(k) \end{bmatrix} \; ; \quad \underline{B}^a(k) = \begin{bmatrix} \underline{B}(k) \\ \cdots \\ \underline{0} \end{bmatrix} \; ;$$

$$\underline{\xi}^a(k) = \begin{bmatrix} \underline{\xi}(k) \\ \cdots \\ \underline{\xi}(k) \end{bmatrix} \tag{5.3.25}$$

$$\underline{C}^a(k) = [\underline{C}(k) \; \vdots \; \underline{I}_n] \quad .$$

The cost (5.3.1) can be written as

$$J^a(\underline{u}) = E\left\{\underline{x}^{a\prime}(N)\underline{F}^a\underline{x}^a(N) + \sum_{k=0}^{N-1}\left(\underline{x}^{a\prime}(k)\underline{W}^a(k)\underline{x}^a(k) + \underline{u}'(k)\underline{M}(k)\underline{u}(k)\right)\right\} \tag{5.3.26}$$

where

$$\underline{F}^a = \begin{bmatrix} \underline{F} & \vdots & \underline{0} \\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{0} \end{bmatrix} \; ; \quad \underline{W}^a(k) = \begin{bmatrix} \underline{W}(k) & \vdots & \underline{0} \\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{0} \end{bmatrix} \quad . \tag{5.3.27}$$

The augmented control problem is to find $\underline{u}(k)$, of the form (5.2.4) such that the augmented cost (5.3.26) is minimized subject to the augmented system (5.3.25) and constraint (5.2.4). We note that the solution for the augmented control problem is the same as that of the original control problem.

Apply theorem 5.2.3 to the augmented control problem, we have

$$\underline{u}^*(k) = -\left(\underline{M}(k) + \underline{B}^{a\prime}(k)\underline{K}^a(k+1)\underline{B}^a(k)\right)^{-1}\underline{B}^{a\prime}(k)\underline{K}^a(k+1)\underline{A}^a(k)\underline{\hat{x}}^a(k|k) \tag{5.3.28}$$

$$\underline{K}^a(k) = \underline{A}_a{}'(k)(\underline{K}^a(k+1)-\underline{K}^a(k+1)\underline{B}^a(k)(\underline{M}(k)+\underline{B}^{a\prime}(k)\underline{K}(k+1)\underline{B}^a(k))^{-1}\underline{B}^{a\prime}(k)$$

$$\underline{K}^a(k+1)\underline{A}^a(k)+\underline{W}^a(k) \quad ; \quad \underline{K}^a(N) = \underline{F}^a \qquad (5.3.29)$$

and $\underline{\hat{x}}^a(k|k)$ is given by:

$$\mathcal{E}_T^a: \begin{cases} \underline{z}(k+1) = \underline{T}(k+1)\underline{A}^a(k)\underline{P}(k)\underline{z}(k) + \underline{T}(k+1)\underline{A}^a(k)\underline{V}^*(k)\underline{y}(k) + \underline{T}(k+1)\underline{B}^a(k)\underline{u}^*(k) \\ \underline{\hat{x}}^a(k|k) = \underline{P}(k)\underline{z}(k) + \underline{V}^*(k)\underline{y}(k) \quad ; \quad \underline{z}(0) = \underline{T}(0)\underline{x}_o^a \end{cases}$$

(5.3.30)

$$\underline{V}^*(0) = \begin{bmatrix} \underline{\Sigma}_o C'(0) \\ \cdots \\ \underline{\Sigma}_o^n \end{bmatrix} \{\underline{C}(0)\underline{\Sigma}_o\underline{C}'(0) + \underline{\Sigma}_o^n\}^{-1} \qquad (5.3.31)$$

$$\underline{\Sigma}^a(k+1) = \underline{\Delta}^a(k) - \underline{V}(k+1)\underline{C}^a(k+1)\underline{\Delta}^a(k) \qquad (5.3.32)$$

$$\underline{\Delta}^a(k) = \underline{A}^a(k)\underline{\Sigma}^a(k)\underline{A}^{a\prime}(k) + \underline{R}^a(k) \quad ; \quad \underline{V}^*(k+1) \in \mathcal{U}_k(\underline{\Sigma}^a(k)) \quad . \qquad (5.3.33)$$

<u>Lemma 5.3.3</u>: The solution of (5.3.29) is given by

$$\underline{K}^a(k) = \begin{bmatrix} \underline{K}(k) & \vdots & \underline{0} \\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{0} \end{bmatrix} \qquad (5.3.34)$$

with $\underline{K}(k)$ given by (5.3.6).

<u>Proof</u>: We shall use the induction method. At $k = N$, (5.3.27), (5.3.6), and (5.3.29) give

$$\underline{K}^a(N) = \begin{bmatrix} \underline{F} & \vdots & \underline{0} \\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{0} \end{bmatrix} = \begin{bmatrix} \underline{K}(N) & \vdots & \underline{0} \\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{0} \end{bmatrix} \quad . \qquad (5.3.35)$$

Assume that the statement is true at k + 1, we have from (5.3.25), (5.3.27), and (5.3.29) that

$$\underline{K}^a(k) = \begin{bmatrix} \underline{A}(k)(\underline{K}(k+1)-\underline{K}(k+1)\underline{B}(k)(\underline{M}(k)+\underline{B}'(k)\underline{K}(k+1)\underline{B}(k))^{-1}\underline{B}'(k)\underline{K}(k+1))\underline{A}(k)+\underline{W}(k) & \vdots & \underline{0} \\ \cdots & & \\ \underline{0} & & \underline{0} \end{bmatrix}$$

$$= \begin{bmatrix} \underline{K}(k) & \vdots & \underline{0} \\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{0} \end{bmatrix} \tag{5.3.36}$$

and so the lemma follows.

Theorem 5.3.4: The control law,

$$\underline{u}^*(k) = -(\underline{M}(k) + \underline{B}'(k)\underline{K}(k+1)\underline{B}(k))^{-1}\underline{B}'(k)\underline{K}(k+1)\underline{A}(k)\hat{\underline{x}}(k|k) \tag{5.3.37}$$

with $\underline{K}(k)$ given by (5.3.6) and

$$\hat{\underline{x}}(k|k) = [\underline{I}_n \vdots \underline{0}_{nm}]\hat{\underline{x}}^a(k|k) \tag{5.3.38}$$

is the optimal control law which minimizes the cost (5.3.1) subject to (5.3.22), (5.3.23), and (5.2.4). The optimal cost to go is

$$J_*(k,\hat{\underline{x}}) = \hat{\underline{x}}'\underline{K}(k)\hat{\underline{x}}+tr\sum_{i=k}^{N-1}[(\underline{\Sigma}^a(i)-\underline{\Sigma}^a(i+1))\underline{K}^a(i+1))+\underline{W}^a(i)\underline{\Sigma}^a(i)]+\underline{F}^a\underline{\Sigma}^a(N) \quad . \tag{5.3.39}$$

This follows easily from theorem 5.3.2, lemma 5.3.3, and equation (5.3.28).

Note that $\underline{x}^a(k) \varepsilon R^{n+m}$, and so $\underline{z}(k) \varepsilon R^n$ (see chapter 3). The detail structure of optimal control system is described in detail in Figure 5.2.

5.4 Estimation with Feedback for Continuous Linear Systems

Consider a continuous linear system S_3^c described by:

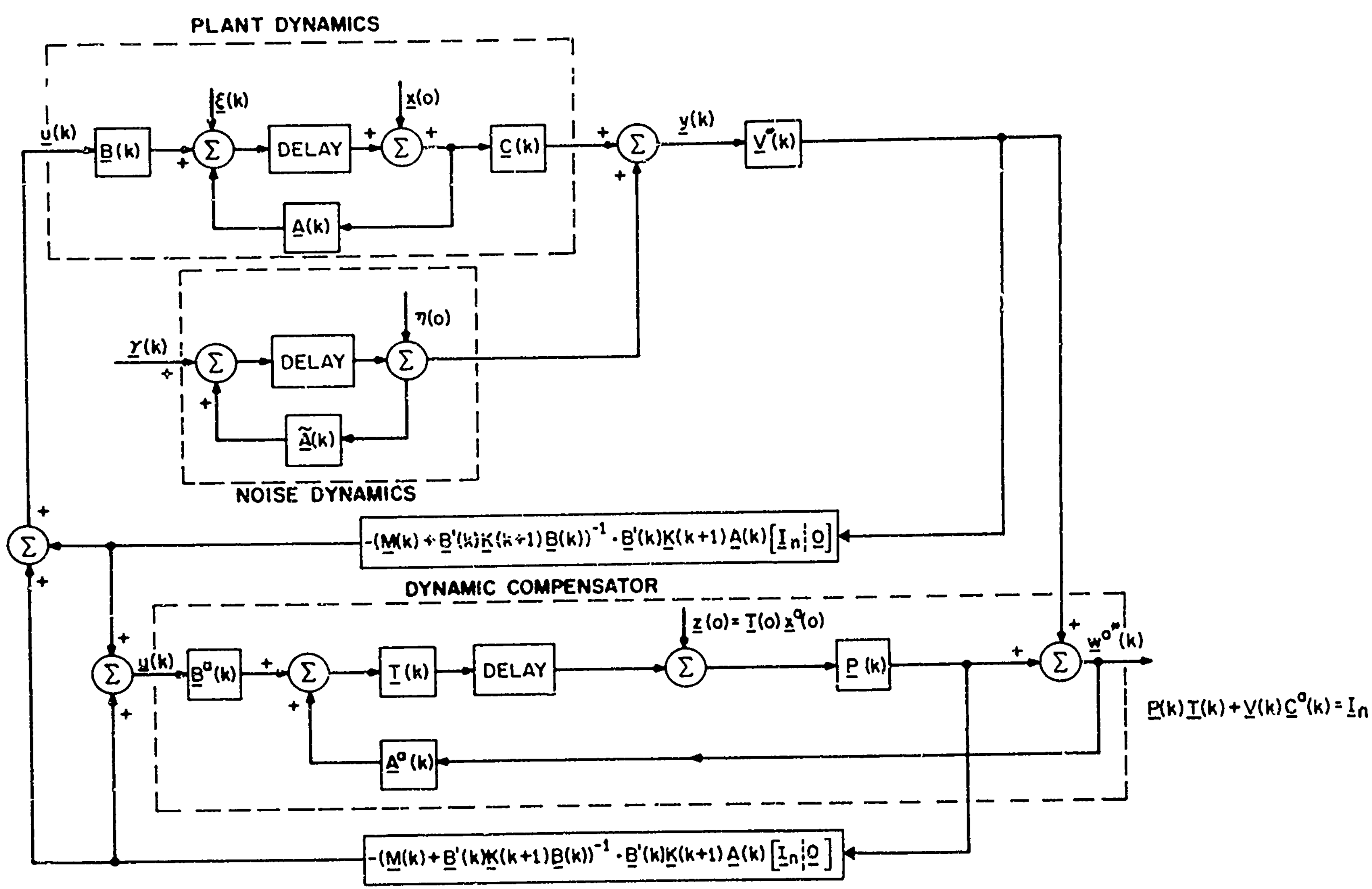


Fig. 5.2 DISCRETE TIME STOCHASTIC LINEAR OPTIMAL CONTROL SYSTEM. THE OBSERVATION NOISE IS SEQUENTIALLY CORRELATED.

$$\mathcal{S}_3^c: \quad \begin{aligned} &\underline{\dot{x}}(t) = \underline{A}(t)\underline{x}(t) + \underline{B}(t)\underline{u}(t) + \underline{\xi}(t) \\ &\underline{y}(t) = \begin{bmatrix} \underline{y}_1(t) \\ \cdots \\ \underline{y}_2(t) \end{bmatrix} = \begin{bmatrix} \underline{C}_1(t)\underline{x}(t) + \underline{\eta}(t) \\ \cdots \\ \underline{C}_2(t)\underline{x}(t) \end{bmatrix} \end{aligned} \tag{5.4.1}$$

where $\underline{x}(t) \in R^n$, $\underline{u}(t) \in R^r$, $\underline{\xi}(t) \in R^n$, $\underline{\eta}(t) \in R^{m_1}$, $\underline{y}(t) \in R^m$. We assume that $\underline{x}(t_o)$, $\{\underline{\xi}(t),\ t \geq t_o\}$, $\{\underline{\eta}(t),\ t \geq t_o\}$ are independent statistics. $\underline{x}(t_o) \sim G(\underline{x}_o, \underline{\Sigma}_o)$ and $\underline{\xi}(t)$, $\underline{\eta}(t)$, $t \geq t_o$ are white Gaussian noises with properties (4.3.2). The control $\underline{u}(t)$ is feedback in nature.

Let us denote $U[0,t) = \{\underline{u}(\tau):\tau \in [0,t)\}$, and $Y_{U[0,t)}[0,t] = \{\underline{y}_{U[0,t)}(\tau):\tau \in [0,t]\}$. The observation statistic at time t is $\underline{y}_{U[0,t)}(t)$ (the subscript is to indicate that the statistic is dependent on the past control values). The accumulative observation statistic at time t is $Y_{U[0,t)}[0,t]$. We shall assume that the control at time t is a function of accumulative observation statistic:

$$\underline{u}(t) = \underline{\phi}(t, Y_{U[0,t)}[0,t]) \quad . \tag{5.4.2}$$

Denote $F(t,U[0,t)) = F(Y_{U[0,t)}[0,t])$. The control $\underline{u}(t)$ is a random vector which is $F(t,U[0,t))$-measurable.

Let $\underline{f}(s)$ be continuous on $[0,t]$ with values in R^m, define the extension of $\underline{f}(s)$ by

$$(\pi_t \underline{f})(s) = \begin{cases} \underline{f}(s) & 0 \leq s \leq t \\ \underline{f}(t) & t \leq s \leq T \end{cases} \tag{5.4.3}$$

$\pi_t \underline{f}$ thus defined is in $C_m[0,T]$, the class of continuous function defined on $[0,T]$ with values in R^m. The control (5.4.2) can be expressed as

$$\underline{u}(t) = \underline{\phi}(t, -_t Y_{U[0,t)}[0,t]) \tag{5.4.4}$$

where $\underline{\phi}(\cdot,\cdot)$ is viewed as a mapping from $R \times C_m[0,T] \to R^r$. The control (5.4.4) is also $F(t,U[0,t))$-measurable. We assume that $\underline{\phi}(t,\cdot)$ satisfies a Lipschitz condition:

$$\|\underline{\phi}(t,f) - \underline{\phi}(t,g)\| < \alpha_1 \|f - g\| \quad ; \quad f,g \in C_m[0,T] \tag{5.4.5}$$

for all $t \in [0,T]$ where α_1 is some constant. Sometimes we shall subpress the dependence on past control value, and write (5.4.4) as

$$\underline{u}(t) = \underline{\phi}(t, -_t y) \quad ; \quad y \triangleq \{\underline{y}(\tau): \tau \in [0,t]\} \tag{5.4.6}$$

without causing confusion.

<u>Theorem 5.4.1</u>: Let $\underline{C}_2(t)\underline{R}(t)\underline{C}_2'(t) > \underline{0}$, and the control is of the feedback form (5.4.6). The conditional distribution of the current state of S_3^c is Gaussian random vector, and is parameterized by the conditional mean, $\hat{\underline{x}}(t|t)$, and conditional covariance, $\underline{\Sigma}(t)$, which are given by:

$$\begin{aligned}
\mathcal{E}_T^{3c}(L^*):\ \dot{\underline{z}}(t) &= (\underline{T}(t)\underline{A}(t)\underline{P}(t)+\dot{\underline{T}}(t)\underline{P}(t)-\underline{T}(t)\underline{L}_1^*(t)\underline{C}_1(t)\underline{P}(t))\underline{z}(t)+\underline{T}(t)\underline{L}_1^*(t)\underline{y}_1(t)\\
&\quad+(\underline{T}(t)\underline{A}(t)\underline{V}_2^*(t)+\dot{\underline{T}}(t)\underline{V}_2^*(t)-\underline{T}(t)\underline{L}_1^*(t)\underline{C}_1(t)\underline{V}_2^*(t))\underline{y}_2(t)\\
&\quad+\underline{T}(t)\underline{B}(t)\underline{u}(t)\\
\underline{z}(t_o^+) &= \underline{T}(t_o)\underline{x}_o-\underline{T}(t_o)\underline{V}_2^*(t_o^+)\underline{y}_2(t_o)\\
\hat{\underline{x}}(t|t) &= \underline{P}(t)\underline{z}(t)+\underline{V}_2^*(t)\underline{y}_2(t)
\end{aligned} \tag{5.4.7}$$

$$\begin{aligned}
\dot{\underline{\Sigma}}(t) &= (\underline{A}(t)-\underline{R}(t)\underline{C}_2'(t)\underline{\Delta}^{-1}(t)\tilde{\underline{C}}_2(t))\underline{\Sigma}(t)+\underline{\Sigma}(t)(\underline{A}(t)-\underline{R}(t)\underline{C}_2'(t)\underline{\Delta}^{-1}(t)\tilde{\underline{C}}_2(t))'\\
&\quad-\underline{\Sigma}(t)(\tilde{\underline{C}}_2'(t)\underline{\Delta}^{-1}(t)\tilde{\underline{C}}_2(t)+\underline{C}_1'(t)\underline{Q}^{-1}(t)\underline{C}_1(t))\underline{\Sigma}(t)+\underline{R}(t)\\
&\quad-\underline{R}(t)\underline{C}_2'(t)\underline{\Delta}^{-1}(t)\underline{C}_2(t)\underline{R}(t)\\
\underline{\Sigma}(t_o) &= \underline{\Sigma}_o-\underline{\Sigma}_o\underline{C}_2'(t_o)(\underline{C}_2(t_o)\underline{\Sigma}_o\underline{C}_2'(t_o))^{-1}\underline{C}_2(t_o)\underline{\Sigma}_o \quad ; \quad \underline{\Delta}(t) \triangleq \underline{C}_2(t)\underline{R}(t)\underline{C}_2'(t)
\end{aligned} \tag{5.4.8}$$

where $\underline{V}_2^*(t)$, $\underline{L}_1^*(t)$ are given by

$$\underline{V}_2^*(t) = \begin{cases} \underline{\Sigma}_o \underline{C}_2'(t_o)(\underline{\Sigma}_o \underline{C}_2'(t_o))^{-1} & ; \quad ; \quad t = t_o \\ (\underline{\Sigma}(t)\tilde{\underline{C}}_2'(t) + \underline{R}(t)\underline{C}_2'(t))\underline{\mathcal{L}}^{-1}(t) & ; \quad t > t_o \end{cases} \tag{5.4.9}$$

$$\underline{L}_1^*(t) = \underline{\Sigma}(t)\underline{C}_1'(t)\underline{Q}^{-1}(t) \qquad ; \quad t > t_o \tag{5.4.10}$$

and $\underline{P}(t)$, $\underline{T}(t)$ are given by

$$\underline{T}(t)\underline{V}_2^*(t) = \underline{0}_{nm_2} \quad ; \quad \underline{C}_2(t)\underline{P}(t) = \underline{0}_{m_2 n} \quad ; \quad \underline{T}(t)\underline{P}(t) = \underline{I}_{m_2} \qquad t \geq t_o \quad . \tag{5.4.11}$$

<u>Proof</u>: Let us break $\underline{x}(t)$ into

$$\underline{x}(t) = \underline{x}_1(t) + \underline{x}_2(t) \tag{5.4.12}$$

and $\underline{x}_1(t)$, $\underline{x}_2(t)$ are given by

$$\dot{\underline{x}}_1(t) = \underline{A}(t)\underline{x}_1(t) + \underline{B}(t)\underline{u}(t) \quad ; \quad \underline{x}_1(t_o) = \underline{0} \tag{5.4.13}$$

$$\dot{\underline{x}}_2(t) = \underline{A}(t)\underline{x}_2(t) + \underline{\xi}(t) \qquad ; \quad \underline{x}_2(t_o) = \underline{x}(t_o) \tag{5.4.14}$$

$\underline{u}(t)$ is of the form (5.4.6) and is F(t,U[0,t))-measurable; therefore $\underline{x}_1(t)$ is F(t,U[0,t))-measurable. From (5.4.12), we deduce

$$\hat{\underline{x}}(t|t) = \underline{x}_1(t) + E\{\underline{x}_2(t)|F(t,U[0,t))\} \quad . \tag{5.4.15}$$

Let us define

$$\underline{y}_1(t) = \underline{C}(t)\underline{x}_1(t) \quad ; \quad \underline{y}_2(t) = \underline{y}(t) - \underline{y}_1(t) = \underline{C}(t)\underline{x}_2(t) + \begin{bmatrix} \underline{\eta}(t) \\ \cdots \\ \underline{0} \end{bmatrix} \quad . \tag{5.4.16}$$

Let $F_2(t) = F\{\underline{y}_2(\tau), \tau \in [0,t)\}$. Since $\{\underline{y}(s), \underline{y}_1(s)\}_{s=0}^{t}$ are $F(t,U[0,t))$-measurable, then $\{\underline{y}_2(s)\}_{s=0}^{t}$ is $F(t,U[0,t))$-measurable and so

$$F_2(t) \subset F(t,U[0,t)) \quad . \tag{5.4.17}$$

From (5.4.16) and (5.4.13), we have

$$\underline{y}(t) = \underline{y}_2(t) + \underline{C}(t) \int_{t_o}^{t} \underline{\phi}_A(t,\tau)\underline{B}(\tau)\underline{u}(\tau)d\tau \tag{5.4.18}$$

where $\underline{u}(\tau)$ is of the form (5.4.6). Equation (5.4.18) is an integral equation. By the Lipschitz assumption, equation (5.4.18) can be solved by successive approximations to yield a unique $\underline{y} \in C_m[0,T]$.[6] Setting $\underline{y}^{(o)}(t) \equiv \underline{0}$ and

$$\underline{y}^{(\nu)}(t) = \underline{y}_2(t)+\underline{C}(t)\int_{t_o}^{t} \underline{\phi}_A(t,\tau)\underline{B}(\tau)\underline{\hat{\phi}}(\tau,\pi_\tau y^{(\nu-1)})d\tau$$

$$t \in [0,T] \quad ; \quad \nu = 1, 2, \ldots \quad . \tag{5.4.19}$$

Inductively, $\{\underline{y}^{(\nu)}(s)\}_{s=0}^{t}$ is $F_2(t)$-measurable for $\nu = 1, 2, \ldots$; and so $\{\underline{y}(s)\} = \{\lim_{\nu\to\infty} \underline{y}^{(\nu)}(s)\}_{s=0}^{t}$ is also $F_2(t)$-measurable, and

$$F_2(t) \supset F(t,U[0,t)) \quad . \tag{5.4.20}$$

Combining (5.4.17) and (5.4.20), we have

$$F_2(t) = F(t,U[0,t)) \quad . \tag{5.4.21}$$

Equation (5.4.15) becomes

$$\underline{\hat{x}}(t|t) = \underline{x}_1(t) + E\{\underline{x}_2(t)|F_2(t)\} \quad . \tag{5.4.22}$$

Now consider the systems:

$$\mathcal{S}_1^c: \begin{cases} \dot{\underline{x}}_1(t) = \underline{A}(t)\underline{x}_1(t) + \underline{B}(t)\underline{u}(t) \;;\; \underline{x}_1(t_o) = \underline{0} \\ \underline{y}_1(t) = \underline{C}(t)\underline{x}_1(t) \end{cases} \tag{5.4.23}$$

$$\mathcal{S}_2^c: \begin{cases} \dot{\underline{x}}_2(t) = \underline{A}(t)\underline{x}_2(t) + \underline{\xi}(t) \;; \quad \underline{x}_2(t_o) = \underline{x}(t_o) \sim \mathcal{G}(\underline{x}_o, \underline{\Sigma}_o) \\ \underline{y}_2(t) = \begin{bmatrix} \underline{C}(t)\underline{x}_2(t) + \underline{n}(t) \\ \cdots \\ \underline{C}_2(t)\underline{x}_2(t) \end{bmatrix} \end{cases} \tag{5.4.24}$$

Apply observers theory to the deterministic system $\mathcal{S}_1^c$ and stochastic systems $\mathcal{S}_2^c$ (see chapter 4, sections 4.2, 4.3). In this manner we prove the theorem easily. For detail procedures the reader is referred to theorem 5.2.1, where we have proved the discrete analog in great details.

5.5 Stochastic Control of Continuous Linear Systems with Quadratic Criteria

We consider the problem of controlling the continuous linear system $\mathcal{S}_3^c$ with quadratic criteria

$$J^c(\underline{u}) = E\left\{\underline{x}'(T)\underline{F}\,\underline{x}(T) + \int_{t_o}^{T} \underline{x}'(t)\underline{W}(t)\underline{x}(t) + \underline{u}'(t)\underline{M}(t)\underline{u}(t)dt\right\} \tag{5.5.1}$$

with $\underline{F} \geq \underline{0}$, $\underline{W}(t) \geq \underline{0}$ and $\underline{M}(t) > \underline{0}$. We are to find a control of the form (5.4.4) and (5.4.5) such that (5.5.1) is minimized subject to (5.4.1).

For any control of the form (5.4.4) and (5.4.5) we have from lemma 2.2.6 that

$$J^c(\underline{u}) = E\left\{E\{\underline{x}'(T)\underline{F}\,\underline{x}(T)\,|\,F(T,U[0,T])\} + \int_{t_o}^{T} E\{\underline{x}'(t)\underline{W}(t)\underline{x}(t) + u'(t)\underline{M}(t)\underline{u}(t)\,|\,F(t,U[0,t))\}dt\right\}$$

$$= E\left\{\underline{\hat{x}}'(T|T)\underline{F}\,\underline{\hat{x}}(T|T)+\int_{t}^{T}\underline{\hat{x}}'(t|t)\underline{W}(t)\underline{x}(t|t)+\underline{u}'(t)\underline{M}(t)\underline{u}(t)dt\right\}$$

$$+tr\left\{\underline{F}\,\underline{\Sigma}(T|T)+\int_{t_o}^{T}\underline{W}(t)\underline{\Sigma}(t)dt\right\} \tag{5.5.2}$$

where $\underline{\Sigma}(t)$, $t \geq t_o$, is given by (5.4.8). We note from (5.4.8) that $\underline{\Sigma}(t)$ is independent of the control function; thus to minimize (5.5.1) is equivalent to minimizing

$$J^{c\prime}(\underline{u}) = E\left\{\underline{\hat{x}}'(T|T)\underline{F}\,\underline{\hat{x}}(T|T)+\int_{t_o}^{T}\underline{\hat{x}}'(t|t)\underline{W}(t)\underline{\hat{x}}(t|t)+\underline{u}'(t)\underline{M}(t)\underline{u}(t)dt\right\} \quad . \tag{5.5.3}$$

From (5.4.7), we can easily derive the differential equation for $\underline{\hat{x}}(t|t)$:

$$\dot{\underline{\hat{x}}}(t|t) = \underline{A}(t)\underline{\hat{x}}(t|t)+(\underline{V}_2^*(t)\tilde{\underline{C}}_2(t)-\underline{L}_1^*(t)\underline{C}_1(t))(\underline{\hat{x}}(t|t)-\underline{x}(t))+\underline{V}_2^*(t)\underline{C}_2(t)\underline{\xi}(t)$$

$$+\underline{L}_1^*(t)\underline{n}(t)+\underline{B}(t)\underline{u}(t) \tag{5.5.4}$$

with $\tilde{\underline{C}}_2(t) = \dot{\underline{C}}_2(t) + \underline{C}_2(t)\underline{A}(t)$, and $\underline{x}(t)$, $t \geq t_o$, is a diffusion process given by (5.4.1). We have now a stochastic control problem: <u>Find a control law of the form (5.4.4) and (5.4.5), such that the cost (5.5.3) is minimized subject to the constraints (5.5.4), (5.4.1).</u>

<u>Lemma 5.5.1</u>: The control law

$$\underline{u}^*(t) = -\underline{M}^{-1}(t)\underline{B}'(t)\underline{K}(t)\underline{\hat{x}}(t|t) \tag{5.5.5}$$

$$-\dot{\underline{K}}(t) = \underline{A}'(t)\underline{K}(t)+\underline{K}(t)\underline{A}(t)-\underline{K}(t)\underline{B}(t)\underline{M}^{-1}(t)\underline{B}'(t)\underline{K}(t)+\underline{W}(t) \quad ; \quad \underline{K}(T) = \underline{F} \tag{5.5.6}$$

is the optimal law for the above stochastic control problem, i.e., if $\underline{u}^o(t)$ is a control of the form (5.4.4) and (5.4.5), then

$$J^{c\prime}(\underline{u}^*) \leq J^{c\prime}(\underline{u}^o) \quad . \tag{5.5.7}$$

The optimal cost-to-go is

$$J_{*}^{c\prime}(t,\underline{\hat{x}}) = E\Big\{\underline{\hat{x}}^{*\prime}(T|T)\underline{F}\,\underline{\hat{x}}^{*}(T|T)+\int_t^T \underline{\hat{x}}^{*\prime}(\tau|\tau)\underline{W}(\tau)\underline{\hat{x}}^{*}(\tau|\tau)+\underline{u}^{*\prime}(\tau)\underline{M}(\tau)\underline{u}^{*}(\tau)d\tau$$

$$\underline{\hat{x}}^{*}(t|t) = \underline{\hat{x}}\Big\}$$

$$= \underline{\hat{x}}'\underline{K}(t)\underline{\hat{x}}+\mathrm{tr}\int_t^T (\underline{V}_2^{*}(\tau)\underline{C}_2(\tau)\underline{R}(\tau)\underline{C}_2'(\tau)\underline{V}_2^{*\prime}(\tau)$$

$$+\underline{L}_1^{*}(\tau)\underline{Q}(\tau)\underline{L}_1^{*\prime}(\tau))\underline{K}(\tau)d\tau \quad . \tag{5.5.8}$$

Proof: As in the discrete analog, we shall make use of the Optimality Criterion (theorem 2.4.4) to prove the lemma. Let us define for all $(t,\underline{\hat{x}}) \in [0,T] \times R^n$:

$$C(t,\underline{\hat{x}}) = \underline{\hat{x}}'\underline{K}(t)\underline{\hat{x}}+\mathrm{tr}\left\{\int_t^T (\underline{V}_2^{*}(\tau)\underline{C}_2(\tau)\underline{R}(\tau)\underline{C}_2'(\tau)\underline{V}_2^{*\prime}(\tau)+\underline{L}_1^{*}(\tau)\underline{Q}(\tau)\underline{L}_1^{*\prime}(\tau))\underline{K}(\tau)d\tau\right\} \tag{5.5.9}$$

where $\underline{K}(t)$ satisfies (5.5.6). From (5.5.6) and (5.5.9), we have

$$C(T,\underline{\hat{x}}) = \underline{\hat{x}}'\underline{F}\,\underline{\hat{x}} \quad . \tag{5.5.10}$$

Let $U[0,t)$ be an arbitrary control function and denote

$$\underline{\hat{x}} = E\{\underline{x}(t)|F(t,U[0,t))\} = \underline{\hat{x}}(t|t) \quad . \tag{5.5.11}$$

Let $\underline{u}^{*}(t)$ be given by (5.5.5) with $\underline{\hat{x}}(t|t)$ replaced by $\underline{\hat{x}}$. Denote the differential generator of $\underline{\hat{x}}(t|t)$ by $\mathcal{L}_u(\cdot)$, we have from (5.5.4), (5.5.5), (5.5.6):

$$E\{\mathcal{L}_{u^*}(C(t,\underline{\hat{x}}))+\underline{\hat{x}}'\underline{W}(t)\underline{\hat{x}}+\underline{u}^{*\prime}(t)\underline{M}(t)\underline{u}^{*}(t)|F(t,U[0,t))\}$$

$$= \mathrm{tr}\{(\underline{V}_2^{*}(t)\underline{C}_2(t)\underline{R}(t)\underline{C}_2'(t)\underline{V}_2^{*\prime}(t)+\underline{L}_1^{*}(t)\underline{Q}(t)\underline{L}_1^{*}(t)\underline{K}(t)\}+2\underline{\hat{x}}'\underline{A}'(t)\underline{K}(t)\underline{\hat{x}}+$$

$$2\underline{u}^{*\prime}(t)\underline{B}'(t)\underline{K}(t)\underline{\hat{x}}+\underline{\hat{x}}'\underline{W}(t)\underline{\hat{x}}+\underline{u}^{*\prime}(t)\underline{M}(t)\underline{u}^{*}(t)$$

$$= \mathrm{tr}\{(\underline{V}_2^{*}(t)\underline{C}_2(t)\underline{R}(t)\underline{C}_2'(t)\underline{V}_2^{*}(t)+\underline{L}_1^{*}(t)\underline{Q}(t)\underline{L}_1^{*\prime}(t))\underline{K}(t)\}-\underline{\hat{x}}\,\underline{\dot{K}}(t)\underline{\hat{x}} \quad . \tag{5.5.12}$$

Combining (5.5.9) and (5.5.12) we have

$$C_t(t,\underline{\hat{x}})+E\{\mathcal{L}_{u^*}(C(t,\underline{\hat{x}})+\underline{\hat{x}}'\underline{W}(t)\underline{\hat{x}}+\underline{u}^{*}{}'(t)\underline{M}(t)\underline{u}(t)|F(t,U[0,t))\} = 0 \quad . \quad (5.5.13)$$

Let $\underline{u}^o(t)$ be any $F(t,U[0,t))$-measurable function, we have from (5.5.4), (5.5.5) and (5.5.6)

$$E\{\mathcal{L}_{u^o}(C(t,\underline{\hat{x}})+\underline{\hat{x}}'\underline{W}(t)\underline{\hat{x}}+\underline{u}^{o}{}'(t)\underline{M}(t)\underline{u}^o(t)|F(t,U[0,t))\}$$

$$= tr\{(\underline{V}_2^*(t)\underline{C}_2(t)\underline{R}(t)\underline{C}_2'(t)\underline{V}_2^*{}'(t)+\underline{L}_1^*(t)\underline{Q}(t)\underline{L}_1^*{}'(t))\underline{K}(t)\}-\underline{\hat{x}}'\underline{\hat{K}}(t)\underline{\hat{x}}(t)$$

$$+(\underline{u}^o(t)-\underline{u}^*(t))'\underline{M}(t)(\underline{u}^o(t)-\underline{u}^*(t)) \quad . \quad (5.5.14)$$

Since $\underline{M}(t) > \underline{0}$, (5.5.13) and (5.5.14) imply

$$0 = C_t(t,\underline{\hat{x}})+E\{\mathcal{L}_{u^*}(C(t,\underline{\hat{x}}))+\underline{\hat{x}}'\underline{W}(t)\underline{\hat{x}}+\underline{u}^{*}{}'(t)\underline{M}(t)\underline{u}^*(t)|F(t,U[0,t))\}$$

$$\leq C_t(t,\underline{\hat{x}})+E\{\mathcal{L}_{u^o}(C(t,\underline{\hat{x}}))+\underline{\hat{x}}'\underline{W}(t)\underline{\hat{x}}+\underline{u}^{o}{}'(t)\underline{M}(t)\underline{u}^o(t)|F(t,U[0,t))\} \quad . \quad (5.5.15)$$

The lemma follows from (5.5.10), (5.5.15), (5.5.9) and the Optimality Criterion (theorem 2.4.4).

From lemma 5.5.1 and equation (5.5.2), we have easily the following:

<u>Theorem 5.5.2</u>: The control law $\underline{u}^*(t)$ given by (5.5.5) and (5.5.6) is the optimal control law which minimizes the cost (5.5.1) subject to the constraints (5.4.1), (5.4.4) and (5.4.5). The optimal cost-to-go can be expressed as

$$J_*^c(t,\underline{\hat{x}}) = \underline{\hat{x}}'\underline{K}(t)\underline{\hat{x}}+tr\,\underline{F}\,\underline{\Sigma}(T)+\int_{t_o}^{T}[\underline{W}(\tau)\underline{\Sigma}(\tau)+\underline{V}_2^*(\tau)\underline{C}_2(\tau)\underline{R}(\tau)\underline{C}_2'(\tau)\underline{V}_2^*{}'(\tau)\underline{K}(\tau)$$
$$+\underline{L}_1^*(\tau)\underline{Q}(\tau)\underline{L}_1^*(\tau)\underline{K}(\tau)]d\tau \quad (5.5.16)$$

where $\underline{L}(\cdot)$, $\underline{V}_2^*(\cdot)$, $\underline{L}_1^*(\cdot)$ are given by (5.4.8) to (5.4.10), and $\underline{K}(t)$ is given by (5.5.6).

The structure of the optimal control system for S_3^c is described in Figure 5.3, where we have decomposed the control law into

$$\underline{u}^*(t) = \underline{u}_1^*(t) + \underline{u}_2^*(t) \tag{5.5.17}$$

$\underline{u}_1^*(t)$ is the pure feedback from the noise-free observation:

$$\underline{u}_1^*(t) = -\underline{M}^{-1}(t)\underline{B}'(t)\underline{K}(t)\underline{V}_2^*(t)\underline{y}_2(t) \tag{5.5.18}$$

and $\underline{u}_2^*(t)$ is a feedback after compensation:

$$\underline{u}_2^*(t) = -\underline{M}^{-1}(t)\underline{B}'(t)\underline{K}(t)\underline{P}(t)\underline{z}(t) \quad . \tag{5.5.19}$$

In the special case when $\underline{C}_2(t) = \underline{0}$, i.e., all observation is noisy, we have the usual separation results due to Wonham [27].

The general results can be applied to the case where we have time-correlated observation noise.

Consider the system S_3^c described by (4.5.52), the statistical law of underlying certainties are given by (4.5.53) and (4.5.34). From these assumptions we can form the augmented system S_a^c given by (4.5.55)-(4.5.57). Let us define

$$\underline{W}^a(t) = \begin{bmatrix} \underline{W}(t) & \vdots & \underline{0} \\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{0} \end{bmatrix} \quad ; \quad \underline{F}^a = \begin{bmatrix} \underline{F} & \vdots & \underline{0} \\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{0} \end{bmatrix} \quad . \tag{5.4.20}$$

We form the augmented cost

$$J_a^c(\underline{u}) = E\left\{\underline{x}^{a\prime}(T)\underline{F}^a\underline{x}^a(T)+\int_{t_o}^{T} \underline{x}^{a\prime}(t)\underline{W}^a(t)\underline{x}^a(t)+\underline{u}'(t)\underline{M}(t)\underline{u}(t)\,dt\right\} \quad . \tag{5.4.21}$$

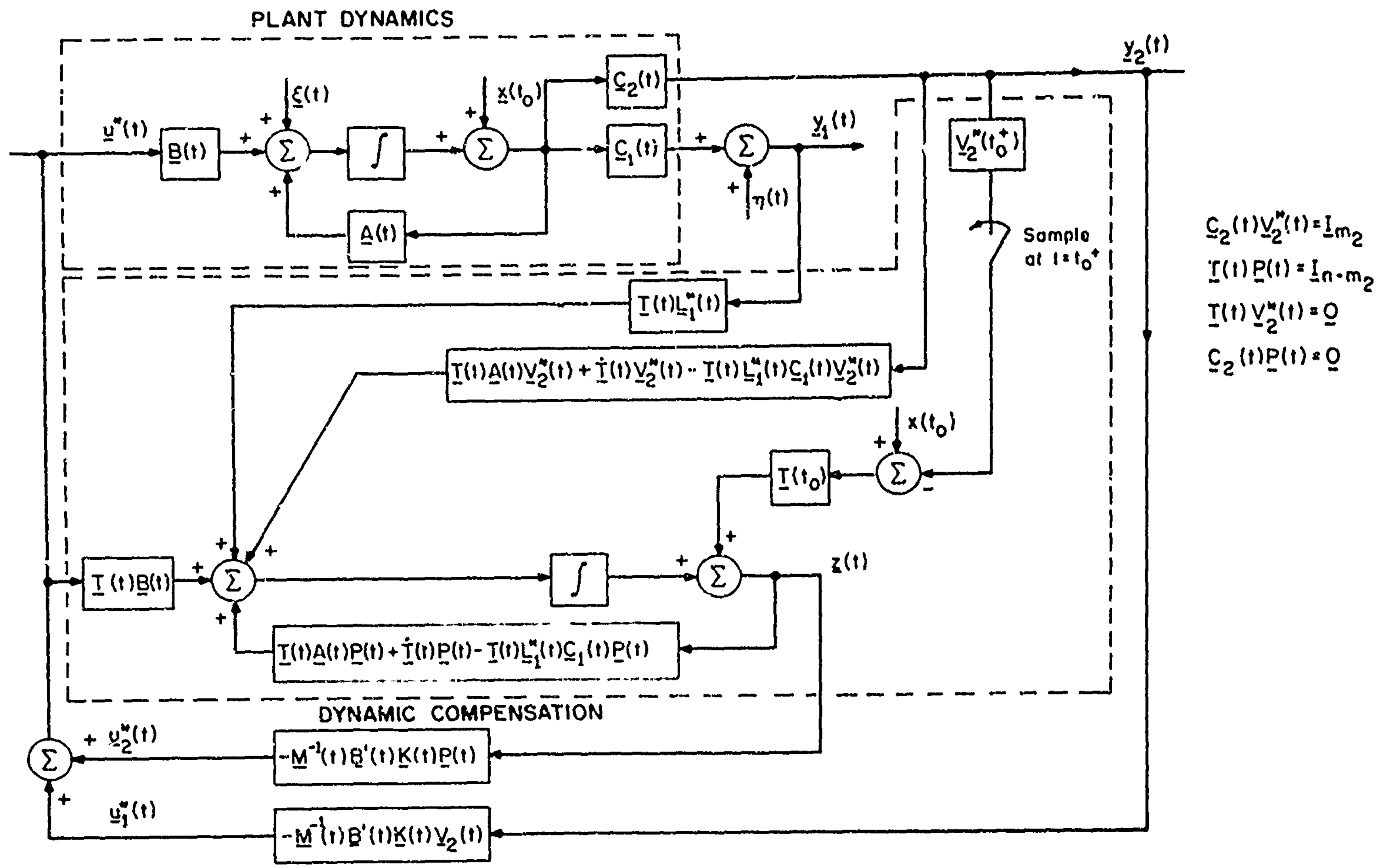


Fig. 5.3 CONTINUOUS TIME STOCHASTIC LINEAR OPTIMAL CONTROL SYSTEM. SOME OBSERVATION CHANNELS ARE NOISE-FREE.

The augmented control problem is to find a control of the form (5.4.4), (5.4.5) which will minimize (5.4.21) subject to the constraint (4.3.53). Note that the solution for the augmented control problem is the same solution for the original control problem where we are to find control of form (5.4.4), (5.4.5) so as to minimize (5.5.1) subject to the dynamical system .

Apply theorem 5.5.2 to the augmented control problem, we have the optimal control law given by

$$\underline{u}^*(t) = -\underline{M}^{-1}(t)\underline{B}^{a\prime}(t)\underline{K}^a(t)\hat{\underline{x}}^a(t|t) \tag{5.4.22}$$

$$-\dot{\underline{K}}^a(t) = \underline{A}^{a\prime}(t)\underline{K}^a(t)+\underline{K}^a(t)\underline{A}^a(t)-\underline{K}^a(t)\underline{B}^a(t)\underline{M}^{-1}(t)\underline{B}^{a\prime}(t)\underline{K}^a(t)+\underline{W}^a(t) \quad ;$$
$$\underline{K}^a(T) = \underline{F}^a \tag{5.4.23}$$

and $\hat{\underline{x}}^a(t|t)$ the conditional mean of $\underline{x}^a(t)$, and is generated via a minimal order optimum observer-estimator (see theorem 5.4.1).

Lemma 5.5.3: The solution of (5.4.23) is

$$\underline{K}^a(t) = \begin{bmatrix} \underline{K}(t) & \vdots & \underline{0} \\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{0} \end{bmatrix} \tag{5.4.24}$$

with $\underline{K}(t)$ satisfying (5.5.6).

Proof: Partition $\underline{K}^a(t)$ into

$$\underline{K}^a(t) = \begin{bmatrix} \underline{K}_{11}(t) & \vdots & \underline{K}_{12}(t) \\ \cdots & & \cdots \\ \underline{K}_{21}(t) & \vdots & \underline{K}_{22}(t) \end{bmatrix} \tag{5.4.25}$$

(5.4.23) gives:

$$-\dot{\underline{K}}_{11}(t) = \underline{A}'(t)\underline{K}_{11}(t)+\underline{K}_{11}(t)\underline{A}(t)-\underline{K}_{11}(t)\underline{B}(t)\underline{M}^{-1}(t)\underline{B}'(t)\underline{K}_{11}(t)+\underline{W}(t) \quad ; \quad \underline{K}_{11}(T) = \underline{F}$$

$$-\dot{\underline{K}}_{12}(t) = \underline{A}'(t)\underline{K}_{12}(t)+\underline{K}_{12}(t)\hat{\underline{A}}(t)-\underline{K}_{11}(t)\underline{B}(t)\underline{M}^{-1}(t)\underline{B}'(t)\underline{K}_{12}(t) \quad ; \quad \underline{K}_{12}(T) = \underline{0} \tag{5.4.26}$$

$$-\dot{\underline{K}}_{22}(t) = \hat{\underline{A}}'(t)\underline{K}_{22}(t)+\underline{K}_{22}(t)\hat{\underline{A}}(t)-\underline{K}_{21}(t)\underline{B}(t)\underline{M}^{-1}(t)\underline{B}(t)\underline{K}_{12}(t) \quad ; \quad \underline{K}_{22}(T) = \underline{0}$$

$$-\underline{K}_{21}(t) = -\underline{K}'_{12}(t) \quad .$$

Comparing with (5.5.6), we see that

$$\underline{K}_{11}(t) = \underline{K}(t) \quad . \tag{5.4.27}$$

From the second equation of (5.4.26), we deduce

$$\underline{K}_{12}(t) = \underline{0} \tag{5.4.28}$$

substituting (5.4.28) into the third equation of (5.4.26) and then we have

$$\underline{K}_{22}(t) = \underline{0} \quad . \tag{5.4.29}$$

Combining (5.4.25) to (5.4.28), we have (5.4.24).

Using lemma 5.5.3, theorem 5.5.2 and equation (5.4.22), we have the results:

<u>Theorem 5.5.4</u>: The control law

$$\underline{u}^*(t) = -\underline{M}^{-1}(t)\underline{B}'(t)\underline{K}(t)\hat{\underline{x}}(t|t) \tag{5.4.30}$$

with $\underline{K}(t)$ satisfying (5.5.6) and

$$\hat{\underline{x}}(t|t) = [\underline{I}_n \vdots \underline{0}]\hat{\underline{x}}^a(t|t) = E\{\underline{x}(t)|F(t,U^*[0,t])\}; \qquad (5.4.31)$$

is the optimal control law of the form (5.4.4), (5.4.5) which minimizes (5.5.1) subject to the dynamical constraints $\tilde{S}_3^c$, (4.5.28). (See Fig. 5.4.)

5.6 General Discussions

In this chapter, we considered the problem of controlling a linear system with quadratic criteria under the assumptions that

1) System dynamics are known,
2) Statistical laws of underlying uncertainties are known.

It has been shown that under fairly general assumptions on the noise structures, the optimal control strategy can be split into two distinct procedures:

1) Find the conditional mean estimates of the current state
2) optimally feedback as if the conditional mean estimate of the current state is the true state of the system.

This result is generally referred to as Separation Theorem[32] or Certainty-equivalence principle.[43] Theorem 5.3.2 includes as special case the results obtained by Joseph and Tou,[56] Gunckel and Franklin;[27] and theorem 5.5.1 generalized that of Wonham's.[27] In the following, we shall discuss some further extensions of the research related to this chapter.

(A) Different Cost Criteria

In this chapter, we have considered exclusively quadratic criteria. The first reason for doing this is motivated by the perturbation guidance approach to many guidance control problem,[43] where we try to keep a stochastic system on a precomputed nominal trajectory. Such an approach will naturally lead to the problem of controlling a time-varying linear

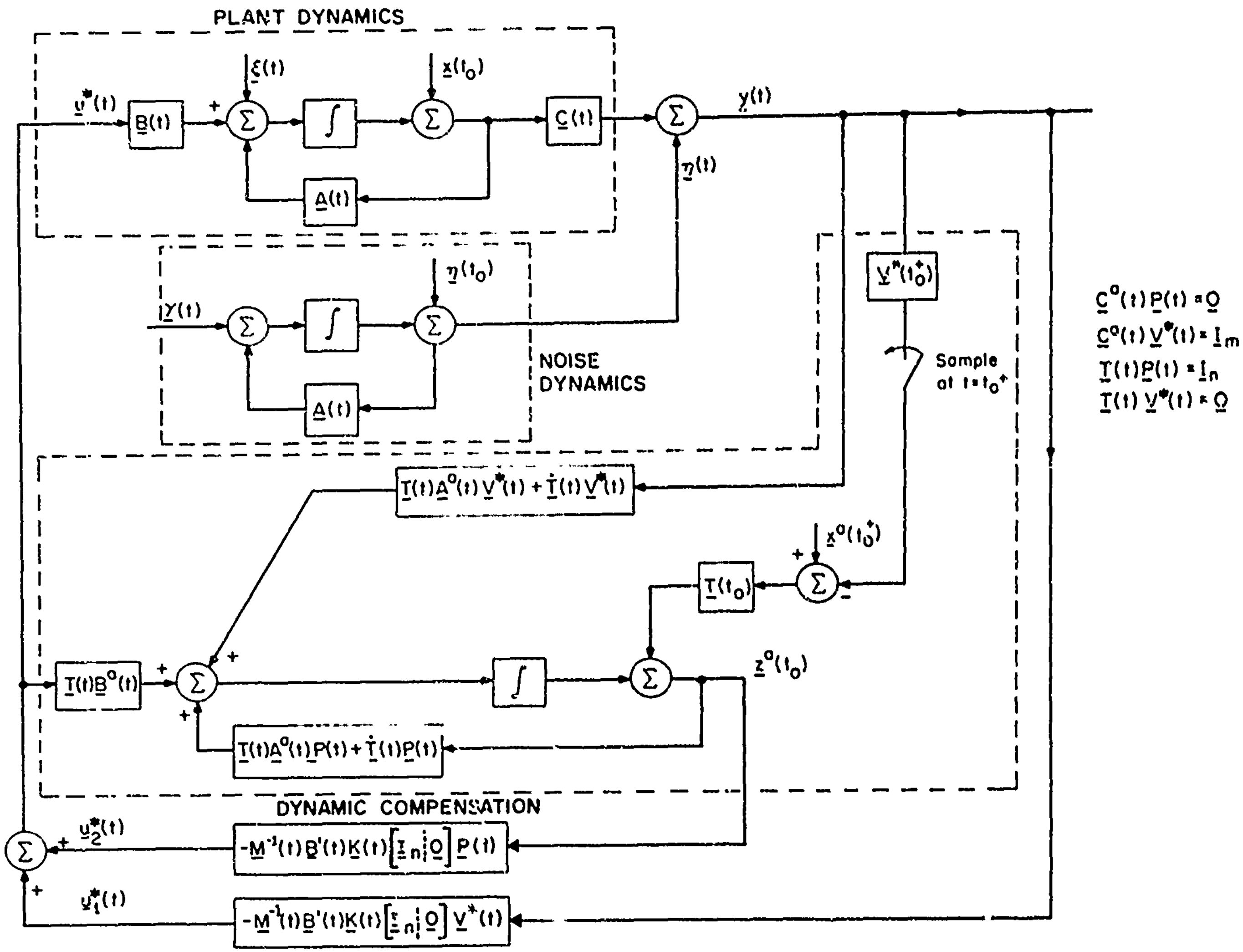


Fig. 5.4 CONTINUOUS TIME STOCHASTIC LINEAR OPTIMAL CONTROL SYSTEM. THE OBSERVATION NOISE IS TIME-CORRELATED.

system with quadratic criteria. There is also the reason that control with quadratic criteria is one special case where we can derive explicit results.

The approach taken in this chapter follows that of Streibel[59] in the discrete case, and that of Wonham[32] in the continuous version. Theoretically, we can easily extend sections 5.3 and 5.5 to more general situations where the cost criteria is not necessary quadratic. The main difficulty that we shall face is the existence problem, which is a mathematical rather than conceptual issue. In general, we shall have to formulate and solve a new stochastic control problem where the process being controlled is the "estimated" process $\underline{\hat{x}}(t|t)$, rather than the process $\underline{x}(t)$. The interested readers are referred to Streibel[59] and Wonham[32] for detail discussions.

(B) Terminal Time $N \to \infty (T \to \infty)$

In the discrete case, let us define $\underline{K}(k,N;\underline{F})$ as:

$$\underline{K}(k,N;\underline{F}) = \underline{A}'(k)(\underline{K}(k+1,N;\underline{F})-\underline{K}(k+1,N;\underline{F})\underline{B}(k)(\underline{M}(k)+\underline{B}'(k)\underline{K}(k+1,N;\underline{F})\underline{B}(k))^{-1}\cdot$$

$$\underline{B}'(k)\underline{K}(k+1,N;\underline{F})\underline{A}(k)+\underline{W}(k) \quad ; \quad \underline{K}(N,N;\underline{F}) = \underline{F} \quad . \qquad (5.6.1)$$

From the separation results, the overall control system can be studied separately by first considering the minimal order optimum observer-estimator, and then the feedback control. In the case when $N \to \infty$, the error covariance will remain bounded if and only if the system S_2 is detectable (see chapter 3). Thus detectability is necessary in order we can reasonably talk about controlling the system during an infinite time span. Next, we have to consider under what appropriate conditions the

feedback gain will remain bounded. We note that this is equivalent to consider under what assumptions will $\underline{K}(k,N;\underline{F})$ remain bounded as $N \to \infty$. Comparing (5.6.1) with (2.5.15) where we replace

$$\begin{aligned} \underline{A}(N-k+k_o) &\to \underline{A}'(k) \\ \underline{0} &\to \underline{Q}_1(k) \\ \underline{W}(N-k+k_o) &\to \underline{Q}_2(k) \\ \underline{M}(N-k+k_o) &\to \underline{R}(k) \\ \underline{B}(N-k+k_o) &\to \underline{A}'(k)\underline{D}'(k) \\ \underline{K}(N-k+k_o,N;\underline{F}) &\to \underline{P}(k,k_o;\underline{F}) \end{aligned} \tag{5.6.2}$$

We can view $\underline{K}(k,N;\underline{F})$ as the minimal sequence with respect to a certain solution set. This allows us to consider the asymptotic behavior of $\underline{K}(k,N;\underline{F})$ as $N \to \infty$. From section 4.6, we see that a necessary and sufficient condition for $\lim_{N\to\infty} \underline{K}(k,N;\underline{F})$ to remain bounded and satisfy a steady-state difference equation is that there exists some matrix $\underline{G}(k)$, $k = \ldots, -1, 0, 1, \ldots$ such that

$$\|\underline{\Phi}_{\hat{A}}(i,j)\| \le \alpha_1 e^{-\alpha_2 |i-j|} \tag{5.6.3}$$

where

$$\underline{\hat{A}}(k) = \underline{A}(k)-\underline{B}(k)\underline{G}(k) \; ; \quad \underline{\Phi}_{\hat{A}}(i,j) = \underline{\hat{A}}(i)\underline{\hat{A}}(i-1)\ldots\underline{\hat{A}}(j) \quad . \tag{5.6.4}$$

Note that (5.6.3) and (5.6.4) are equivalent to saying that there exists $\underline{G}(k)$ such that if we use the control

$$\underline{u}(k) = -\underline{G}(k)\underline{x}(k) \quad . \tag{5.6.5}$$

The resulting system

$$\underline{x}(k+1) = (\underline{A}(k) - \underline{B}(k)\underline{G}(k))\underline{x}(k) + \underline{\xi}(k) \tag{5.6.6}$$

is uniformly asymptotically stable. We shall call such a system <u>stabilizable</u>. Thus <u>detectability</u> and <u>stabilizability</u> are necessary and sufficient conditions which allow us to consider control of discrete linear system over an infinite time span.

In the continuous time case, let us define $\underline{K}(t,T;\underline{F})$ as the solution of

$$-\dot{\underline{K}}(t,T;\underline{F}) = \underline{A}'(t)\underline{K}(t,T;\underline{F})+\underline{K}(t,T;\underline{F})\underline{A}(t)-\underline{K}(t,T;\underline{F})\underline{B}(t)\underline{M}^{-1}(t)\underline{B}'(t)\cdot$$
$$\underline{K}(t,T;\underline{F})+\underline{W}(t) \quad ; \quad \underline{K}(t,T;\underline{F}) = \underline{F} \quad . \tag{5.6.7}$$

In order that we can consider the problem of controlling the continuous linear system $\mathcal{S}_3^c$ during an infinite time span, first we have to require that the error covariance will remain bounded as $T \to \infty$. A sufficient condition for this is detectability of the system $\mathcal{S}_3^c$. Next, we have to consider the asymptotic behavior of $\underline{K}(t,T;\underline{F})$ as $T \to \infty$. Comparing (5.6.7) with (4.3.29) where we replace

$$\begin{aligned}
\underline{A}'(-t) &\rightarrow \underline{A}(t) \\
\underline{0} &\rightarrow \underline{C}_2(t) \\
\underline{B}'(-t) &\rightarrow \underline{C}_1(t) \\
\underline{M}(-t) &\rightarrow \underline{Q}(t) \\
\underline{W}(-t) &\rightarrow \underline{R}(t) \\
\underline{K}(-t,T;\underline{F}) &\rightarrow \underline{\Sigma}^*(t)
\end{aligned} \tag{5.6.8}$$

We have from theorem 4.4.4 that $\lim_{T\to\infty} \underline{K}(t,T;\underline{F})$ will remain bounded if and only if there exists a $\underline{G}(t)$ such that $(\underline{A}(t) - \underline{B}(t)\underline{G}(t))$ is exponentially stable. This is equivalent to the condition that there exists a feedback control

$$\underline{u}(t) = -\underline{G}(t)\underline{x}(t) \tag{5.6.9}$$

such that the resulting system

$$\dot{\underline{x}}(t) = (\underline{A}(t) - \underline{B}(t)\underline{G}(t))\underline{x}(t) + \underline{\xi}(t) \tag{5.6.10}$$

will be uniformly asymptotically stable. We shall call such a system stabilizable. Therefore, in the continuous case, detectability and stabilizability are sufficient conditions which allow us to consider control of continuous linear system over an infinite time span.

With the assumptions on detectability and stabilizability, the asymptotic optimal cost rate is (see (5.3.18))

$$\lim_{N\to\infty} \frac{1}{N-k} J_*(k,\underline{x}) = \lim_{N\to\infty} \frac{1}{N-k}\, \mathrm{tr} \sum_{i=k}^{N-1} (\underline{L}(i) - \underline{\Sigma}(i+1)\underline{K}(i+1) + \underline{W}(i)\underline{\Sigma}(i)) \tag{5.6.11}$$

in the discrete case, and (see (5.4.16))

$$\lim_{T\to\infty} \frac{1}{T-t} J_*^c(t,\underline{x}) = \lim_{T\to\infty} \frac{1}{T-t}\, \mathrm{tr} \int_t^T [\underline{W}(\tau)\underline{\Sigma}(\tau) + \underline{V}_2^*(\tau)\underline{C}_2(\tau)\underline{R}(\tau)\underline{C}_2'(\tau)\underline{V}_2^{*\prime}(\tau)\underline{K}(\tau) + \underline{L}_1^*(\tau)\underline{Q}(\tau)\underline{L}_1^{*\prime}(\tau)\underline{K}(\tau)]d\tau \tag{5.6.12}$$

in the continuous case. We note that the asymptotic optimal cost rate is independent of $\underline{x}$.

In the time invariant case, detectability and stabilizability imply (see chapters 3 and 4)

1) We have a time invariant, minimal order, optimum observer-estimator which generates the conditional mean estimate of the current state.

2) We have a constant feedback gain.

Therefore, the optimum control system is also time invariant where one can write transfer functions for it.

The study on the stochastic stability is a topic for further research.

5.7 Perspective

The Separation Theorem, or certainty-equivalence principle, was stated for discrete linear systems by Joseph and Tou,[56] Gunckel and Franklin,[58] Streibel,[59] and for continuous linear systems by Wonham.[22],[27] The assumption was that the observation noise is non-degenerate white Gaussian process.

The consideration in [56], [58], and [27] is that of quadratic criteria and the approach is straightforward application of the Optimality Criterion. The investigations by Streibel[59] and Wonham[22] include more general cost criteria; the approach taken is that of first finding an equation for the conditional mean of the current state, and then formulate a new optimal control problem where the process being controlled is the conditional mean process; finally, appeal to Optimality Criterion. The approach taken in this chapter is that of Wonham's.[22]. The cost criteria we considered is quadratic, but one can easily extend the results to more general cost criteria. The assumption that the observation noise is a nondegenerate white Gaussian process was relaxed. It was proved that Separation holds when the observation noise is one of the following:

1) regular white Gaussian process

2) degenerate white Gaussian process

3) totally singular situation (i.e., noise-free observation)

4) colored noise (i.e., sequentially correlated or time-correlated)

5) summation of colored and white Gaussian noise.

CHAPTER VI

CONTROL OF DISCRETE TIME LINEAR SYSTEMS WITH UNKNOWN GAIN PARAMETERS

6.1 Introduction

We have considered the control of linear systems with unknown dynamics in the last chapter. Now we shall relax some of the assumptions that all dynamics are known. In many practical control problems, we are confronted with the problem of controlling an unknown linear system. We may have a crude idea about the dimension of the system but the zero and pole locations may not be fully known. In this chapter we shall consider linear systems whose poles are known but whose zeroes are unknown. We shall generalize this to the case of a dynamical system in which the gain vector is unknown. Admittedly, the situation in which we are to control a linear system with unknown gain is rare; however, this research effort is necessary and of importance in guiding our way to the problems of controlling an unknown linear dynamical system.

The structure of this chapter is as follows. In section 6.2, we clearly state the problem under investigation. In section 6.3 we formulate the control problem and state the solution. The approach taken is that of Open-Loop Feedback Optimal (O.L.F.O.) control (see section 6.2). Using the Discrete Matrix Minimum Principle, we derive the O.L.F.O. control sequence in section 6.4. The existence and uniqueness of O.L.F.O. control is studied in detail in section 6.5, and the asymptotic convergence properties of the overall system in section 6.6. Section 6.7 is devoted to the discussion of approaches and of the results. Detailed references are given in section 6.8.

Theoretical results derived in this chapter will be applied to the control of third order linear systems with unknown gain. The computer simulation results and discussions will be treated in the next chapter.

6.2 Problem Statement

Let us consider the discrete linear system

$$\mathcal{S}: \begin{aligned} \underline{x}(k+1) &= \underline{A}(k)\underline{x}(k) + \underline{b}(k)u(k) + \underline{\xi}(k) \\ \underline{y}(k) &= \underline{C}(k)\underline{x}(k) + \underline{\eta}(k) \end{aligned} \tag{6.2.1}$$

where $\underline{x}(k)$, $\underline{\xi}(k) \in R^n$, $\underline{y}(k)$, $\underline{\eta}(k) \in R^m$, $\underline{A}(k)$ is a known nxn matrix, $\underline{C}(k)$ is a known mxn matrix, and u(k) is a scalar control. We assume that the "gain" vector $\underline{b}(k)$ is unknown, but we know that it satisfies the difference equation

$$\underline{b}(k+1) = \underline{G}(k)\underline{b}(k) + \underline{\gamma}(k) \tag{6.2.2}$$

where $\underline{G}(k)$ is a known nxn matrix and $\underline{\gamma}(k) \in R^n$. It is assumed that the vectors $\{\underline{x}(0), \underline{b}(0), \underline{\xi}(k), \underline{\eta}(k), \underline{\gamma}(k); k = 0, 1, \ldots\}$ are independent Gaussian random vectors with known statistical laws:

$$\underline{x}(0) \sim \mathcal{G}(\underline{x}_o, \underline{\Sigma}_{xo}) \tag{6.2.3}$$

$$\underline{b}(0) \sim \mathcal{G}(\underline{b}_o, \underline{\Sigma}_{bo}) \tag{6.2.4}$$

$$\underline{\xi}(k) \sim \mathcal{G}(\underline{0}, \underline{R}(k)) \tag{6.2.5}$$

$$\underline{\eta}(k) \sim \mathcal{G}(\underline{0}, \underline{Q}(k)) \tag{6.2.6}$$

$$\underline{\gamma}(k) \sim \mathcal{G}(\underline{0}, \underline{N}(k)) \tag{6.2.7}$$

with $\underline{\Sigma}_{xo} > \underline{0}$, $\underline{\Sigma}_{bo} > \underline{0}$, $\underline{R}(k) \geq \underline{0}$, $\underline{Q}(k) \geq \underline{0}$, $\underline{N}(k) \geq \underline{0}$.

Our objective is to find a control sequence $\{u(0), \ldots, u(N-1)\}$, such that the cost

$$J(u) = \frac{1}{2} E\left\{\underline{x}'(N)\underline{F}\ \underline{x}(N) + \sum_{k=0}^{N-1} (\underline{x}'(k)\underline{W}(k)\underline{x}(k) + h(k)u^2(k)\right\}$$

is minimized subject to (6.2.1) and (6.2.2). The expectation is taken over all underlying random quantities. We shall assume that $\underline{F}$, and $\underline{W}(k)$ are nonnegative definite symmetric matrices, and that h(k) is a <u>positive</u> scalar for each k.

Depending on the kinds of <u>admissible</u> controls that we are allowed to choose, different formulations of the stochastic optimization problem are possible. In the most general setting, we may assume that the control is a <u>random function</u> of the observed data, i.e., $u(k) = L(u|F(U(0,k-1),k)$ is a conditional probability measure on the control space. If the conditional probability measure is regular, then the control is said to be a <u>mixed</u> control law. If the conditional probability measure is singular (Radon measure), then the control is said to be a <u>pure</u> control law. Unfortunately, little can be done at this level of generality where we consider both mixed and pure control laws.

In the next level of generality, we may confine ourselves to consider only pure control laws to be admissible, i.e., the control at each instant is a <u>fixed</u> function of the observed data; in this case, the resulting control will be a random variable through its dependence on the random observed data. This type restriction of admissible control leads to Bellman's equation [25] whose solution may only be approximated.

Finally, we may restrict ourselves to consider only deterministic open loop controls to be admissible; this essentially means that we ignore the

zero-mean random vectors and assume that the system will behave according to its average behavior. Of course, this may not lead to a good control system, especially whenever the covariances of the disturbances are large. To compensate for this, we shall recompute the open-loop optimal deterministic control after reevaluating the state uncertainty of the system at each and every step (time). A control sequence which is optimal in this manner will be called the open-loop feedback optimal (O.L.F.O.) control.[26],[63] Another interpretation of O.L.F.O. control is the following. Assume that we are to control a system without knowing whether any further observations will be available, or if available, we do not know exactly when the data will be observed. Under this situation the principle of optimality is difficult to apply. One logical, and in some sense optimal approach, is to design an optimal control strategy based on the total information available up to the present time, and continue to use this strategy until new information becomes available; then we change our control strategy accordingly.

In this chapter, we shall look for the O.L.F.O. control. We shall see that such a control sequence is, in some sense, "adaptive" in nature.

6.3 Formulation of Control Problem and its Solution

The present time is indexed by k. Let us assume that the control sequence $U^*(0,k-1) \triangleq \{u^*(0), u^*(1), \ldots, u^*(k-1)\}$ has been applied to the system, and that the observation sequence

$Y_{U^*(0,k-1)}(0,k) \triangleq \{y_{U^*(0,i-1)}(i)\}_{i=0}^{k}$ observed. We would like to find a "future" control sequence $U^*(k,N-1) \triangleq \{u(k), \ldots, u(N-1)\}$ so as to minimize the future cost (cost to go) conditioned on the total available

information at the present time. Let us denote the σ-algebra generated by the observed data $Y_{U^*(0,k-1)}(0,k)$ as $F(k,U^*(0,k-1))$; the symbol $U^*(0,k-1)$ is used to denote that the data is really dependent on the past control history. Our aim now is to find the control sequence $U(k,N-1)$ such that the cost to go

$$J(U(k,N-1);U^*(0,k-1),k) \triangleq \frac{1}{2} E\Big\{\underline{x}'(N)\underline{F}\,\underline{x}(N)+\sum_{j=k}^{N-1}\underline{x}'(j)\underline{W}(j)\underline{x}(j)\,|\,F(k,U^*(0,k-1))\Big\}$$

$$+\frac{1}{2}\sum_{j=k}^{N-1}h(j)u^2(j) \qquad (6.3.1)$$

is minimized subject to the constraints (6.2.1) and (6.2.2). The cost has the simple form (6.3.1) because the future control sequence $U(k,N-1)$ is assumed to be deterministic. (If the future controls were assumed to depend on observed data, we could not take the last term of (6.3.1) outside the expectation operation.) It is now possible to formulate the problem so that deterministic optimization techniques can be applied.

Let us define for $j \geq k$,

$$\hat{\underline{x}}(j|k,U^*(0,k-1)) \triangleq E\{\underline{x}(j)|F(k,U^*(0,k-1))\} \qquad (6.3.2)$$

$$\hat{\underline{b}}(j|k,U^*(0,k-1)) \triangleq E\{\underline{b}(j)|F(k,U^*(0,k-1))\} \qquad (6.3.3)$$

$$\underline{e}_x(j|k,U^*(0,k-1)) \triangleq \hat{\underline{x}}(j|k,U^*(0,k-1)) - \underline{x}(j) \qquad (6.3.4)$$

$$\underline{e}_b(j|k,U^*(0,k-1)) \triangleq \hat{\underline{b}}(j|k,U^*(0,k-1)) - \underline{b}(j) \qquad (6.3.5)$$

We note that $\hat{\underline{x}}(j|k,U^*(0,k-1))$ is $F(k,U^*(0,k-1))$-measurable if $j \geq k$, so for $j \geq k$, we have

$$E\{\underline{x}'(j)\underline{M}\,\underline{x}(j)|F(k,U^*(0,k-1))\} = \hat{\underline{x}}'(j|k,U^*(0,k-1))\underline{M}\,\hat{\underline{x}}(j|k,U^*(0,k-1)) +$$

$$E\{\underline{e}_x'(j|k,U^*(0,k-1))\underline{M}\,\underline{e}_x(j|k,U^*(0,k-1))|F(k,U^*(0,k-1))\} \quad (6.3.6)$$

where $\underline{M}$ is an arbitrary nxn matrix. If we define the state-error second-moment matrix

$$\underline{\Sigma}_x(j|k,U^*(0,k-1)) \triangleq E\{\underline{e}_x(j|k,U^*(0,k-1))\underline{e}_x'(j|k,U^*(0,k-1))|F(k,U^*(0,k-1))\} \quad (6.3.7)$$

then using (6.3.6) and (6.3.7), the conditional cost (6.3.1) can be written as follows

$$J(U(k,N-1);U^*(0,k-1),k) = \frac{1}{2}\,\hat{\underline{x}}'(N|k,U^*(0,k-1))\underline{F}\,\hat{\underline{x}}(N|k,U^*(0,k-1))$$

$$\frac{1}{2}\,\mathrm{tr}\,\underline{F}\,\underline{\Sigma}_x(N|k,U^*(0,k-1)) + \frac{1}{2}\sum_{j=k}^{N-1}\{\hat{\underline{x}}'(j|k,U^*(0,k-1))\underline{W}(j)\underline{x}(j|k,U^*(0,k-1)) +$$

$$\mathrm{tr}\,\underline{W}(j)\underline{\Sigma}_x(j|k,U^*(0,k-1)) + h(j)u^2(j)\} \quad (6.3.8)$$

To complete the formulation, we shall have to derive dynamical equations satisfied by $\hat{\underline{x}}(j|k,U^*(0,k-1))$ and $\underline{\Sigma}_x(j|k,U^*(0,k-1))$.

Since all the noise sequences are assumed to be uncorrelated and white, we have (see chapter 2, section 2.3)

$$E\{\underline{\xi}(j)|F(k,U^*(0,k-1))\} = \underline{0} \;\; ; \;\; E\{\underline{\gamma}(j)|F(k,U^*(0,k-1))\} = \underline{0} \;\; , \;\; j \geq k \quad (6.3.9)$$

The admissible control is assumed to be deterministic; hence, (6.2.1), (6.2.3) and (6.3.9) imply that, for $j \geq k$,

$$\hat{\underline{x}}(j+1|k,U^*(0,k-1)) = \underline{A}(j)\hat{\underline{x}}(j|k,U^*(0,k-1)) + \hat{\underline{b}}(j|k,U^*(0,k-1))u(j) \quad (6.3.10)$$

$$\underline{\hat{b}}(j+1|k,U^*(0,k-1)) = \underline{G}(j)\underline{\hat{b}}(j|k,U^*(0,k-1)) \qquad (6.3.11)$$

with initial condition (at the present time j = k)

$$\begin{aligned}\underline{\hat{x}}(k|k,U^*(0,k-1)) &= E\{\underline{x}(k)|F(k,U^*(0,k-1))\} \quad ; \\ \underline{\hat{b}}(k|k,U^*(0,k-1)) &= E\{b(k)|F(k,U^*(0,k-1))\} \quad .\end{aligned} \qquad (6.3.12)$$

From equations (6.3.2) to (6.3.5), (6.2.1), (6.2.2), (6.3.10), and (6.3.11), we obtain the difference equation for the error vectors for $j \geq k$:

$$\begin{bmatrix} \underline{e}_x(j+1|k,U^*(0,k-1)) \\ \cdots \\ \underline{e}_b(j+1|k,U^*(0,k-1)) \end{bmatrix} = \begin{bmatrix} \underline{A}(j) & \vdots & u(j)\underline{I}_n \\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{G}(j) \end{bmatrix} \begin{bmatrix} \underline{e}_x(j|k,U^*(0,k-1)) \\ \cdots \\ \underline{e}_b(j|k,U^*(0,k-1)) \end{bmatrix} - \begin{bmatrix} \underline{\xi}(j) \\ \cdots \\ \underline{\gamma}(j) \end{bmatrix} \qquad (6.3.13)$$

The initial error at j = k only depends on $\{\underline{\xi}(i), \underline{\gamma}(i)\}$, $i \leq k-1$, and $\{\underline{\eta}(i)\}$, $i \leq k$, and so it is independent of $\{\underline{\xi}(i), \underline{\gamma}(i)\}$, $j \geq k$. Also, since all noises are uncorrelated, zero mean, and white Gaussian, (6.2.6) and (6.2.7) imply that

$$\begin{aligned}E\{\underline{\xi}(j)\underline{\xi}'(j)|F(k,U^*(0,k-1))\} &= \underline{R}(j) \quad ; \\ E\{\underline{\gamma}(j)\underline{\gamma}'(j)|F(k,U^*(0,k-1))\} &= \underline{N}(j) \quad .\end{aligned} \qquad (6.3.14)$$

If we define the second-order moment matrix (for $j \geq k$)

$$\underline{\Sigma}(j|k,U^*(0,k-1)) = E\left\{\begin{bmatrix} \underline{e}_x(j|k,U^*(0,k-1)) \\ \cdots \\ \underline{e}_b(j|k,U^*(0,k-1)) \end{bmatrix} [\underline{e}_x'(j|k,U^*(0,k-1)) \vdots \underline{e}_b'(j|k,U^*(0,k-1))] \;\middle|\; F(k,U^*(0,k-1))\right\} \qquad (6.3.15)$$

Equations (6.3.13), (6.3.14) and the independence of initial error ($j = k$) and of the future noise sequence imply that $\underline{\Sigma}(j|k,U^*(0,k-1))$, $j \geq k$, is generated by (see chapter 2, section 2.3)

$$\underline{\Sigma}(j+1|k,U^*(0,k-1)) = \tilde{\underline{A}}(j,u(j))\underline{\Sigma}(j|k,U^*(0,k-1))\tilde{\underline{A}}'(j,u(j)) + \tilde{\underline{R}}(j) \tag{6.3.16}$$

where

$$\tilde{\underline{A}}(j,u(j)) \triangleq \begin{bmatrix} \underline{A}(j) & \vdots & u(j)\underline{I}_n \\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{G}(j) \end{bmatrix} \quad ; \quad \tilde{\underline{R}}(j) \triangleq \begin{bmatrix} \underline{R}(j) & \vdots & \underline{0} \\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{N}(j) \end{bmatrix} \quad . \tag{6.3.17}$$

The initial condition is given by

$$\underline{\Sigma}(k|k,U^*(0,k-1)) = E\left\{\begin{bmatrix} \underline{e}_x(k|k,U^*(0,k-1)) \\ \cdots \\ \underline{e}_b(k|k,U^*(0,k-1)) \end{bmatrix} [\underline{e}'_x(k|k,U^*(0,k-1)) \vdots \underline{e}'_b(k|k,U^*(0,k-1))] \middle| F(k,U^*(0,k-1))\right\} \tag{6.3.18}$$

From (6.3.12) and (6.3.18), we see that $\underline{x}(k|k,U^*(0,k-1))$ and $\underline{b}(k|k,U^*(0,k-1))$ are the conditional means of $\underline{x}(k)$ and $\underline{b}(k)$, respectively, while $\underline{\Sigma}(k|k,U^*(0,k-1))$ is the conditional covariance matrix of the augmented vector

$$\begin{bmatrix} \underline{x}(k) \\ \cdots \\ \underline{b}(k) \end{bmatrix} .$$

These quantities can be generated by the following identification equations, (6.3.19)-(6.3.23), once the past control $U^*(0,k-1)$ has been chosen:

$$\begin{bmatrix} \underline{\hat{x}}(i+1|i+1,U^*(0,i)) \\ \cdots \\ \underline{\hat{b}}(i+1|i+1,U^*(0,i)) \end{bmatrix} = [\underline{I}_{2n}-\underline{V}^*(i+1|i,U^*(0,i))\underline{\tilde{C}}(i+1)]\underline{\tilde{A}}(i,u^*(i)) \cdot$$

$$\begin{bmatrix} \underline{\hat{x}}(i|i,U^*(0,i-1)) \\ \cdots \\ \underline{\hat{b}}(i|i,U^*(0,i-1)) \end{bmatrix} + \underline{V}^*(i+1|i,U^*(0,i))\underline{y}(i+1) \quad ; \quad i = 0, 1, \ldots, k-1$$

$$\begin{bmatrix} \underline{\hat{x}}(0|0,U^*(0,-1)) \\ \cdots \\ \underline{b}(0|0,U^*(0,-1)) \end{bmatrix} \triangleq \begin{bmatrix} \underline{\hat{x}}(0|0) \\ \cdots \\ \underline{\hat{b}}(0|0) \end{bmatrix} = \begin{bmatrix} \underline{x}_o-\underline{\Sigma}_{xo}\underline{C}'(0)(\underline{C}(0)\underline{\Sigma}_{xo}\underline{C}'(0)+\underline{Q}(0))^{-1}(\underline{C}(0)\underline{x}_o-\underline{y}(0)) \\ \cdots \\ \underline{b}_o \end{bmatrix} \quad (6.3.19)$$

where

$$\underline{\tilde{C}}(i+1) \triangleq [\underline{C}(i+1) \; \vdots \; \underline{0}_{mn}] \quad ; \quad i = 0, 1, \ldots, k-1 \quad (6.3.20)$$

and $\underline{V}^*(i+1|i,U^*(0,i))$, $i = 0, 1, \ldots, k-1$, is a solution of the following equations:

$$\underline{V}^*(i+1|i,U^*(0,i))(\underline{\tilde{C}}(i+1)\underline{\tilde{\Delta}}(i|i,U^*(0,i))\underline{\tilde{C}}'(i+1)+\underline{Q}(i+1)) = \underline{\tilde{\Delta}}(i|i,U^*(0,i))\underline{\tilde{C}}'(i+1)$$

$$; \quad i = 0, 1, \ldots, k-1 \quad (6.3.21)$$

$$\underline{\tilde{\Delta}}(i|i,U^*(0,i)) = \underline{\tilde{A}}(i,u^*(i))\underline{\Sigma}(i|i,U^*(0,i-1))\underline{\tilde{A}}'(i,u^*(i)) + \underline{\tilde{R}}(i)$$

$$i = 0, 1, \ldots, k-1 \quad (6.3.22)$$

$$\underline{\Sigma}(i+1|i+1,U^*(0,i)) = \underline{\tilde{\Delta}}(i|i,U^*(0,i))-\underline{V}^*(i+1|i,U^*(0,i))\underline{\tilde{C}}(i+1)\underline{\tilde{\Delta}}(i|i,U^*(0,i))$$

$$i = 0, 1, \ldots, k-1 \quad (6.3.23)$$

$$\underline{\Sigma}(0|0,U^*(0,-1) \triangleq \underline{\Sigma}(0|0) = \begin{bmatrix} \underline{\Sigma}_{xo}-\underline{\Sigma}_{xo}\underline{C}'(0)(\underline{C}(0)\underline{\Sigma}_{xo}\underline{C}'(0)+\underline{Q}(0))^{-1}\underline{C}(0)\underline{\Sigma}_{xo} & \vdots & \underline{0} \\ \cdots & & \cdots \\ \underline{0} & & \underline{\Sigma}_{bo} \end{bmatrix}$$

Referring to chapter 3, section 3.3 we note that the identification equations represent an optimum observer-estimator for the augmented system:

$$\tilde{\mathcal{S}}: \quad \begin{aligned} \begin{bmatrix} \underline{x}(i+1) \\ \underline{b}(i+1) \end{bmatrix} &= \tilde{\underline{A}}(i,u^*(i)) \begin{bmatrix} \underline{x}(i) \\ \underline{b}(i) \end{bmatrix} + \begin{bmatrix} \underline{\xi}(i) \\ \underline{\gamma}(i) \end{bmatrix} \\ \underline{y}(i) &= \tilde{\underline{C}}(i) \begin{bmatrix} \underline{x}(i) \\ \underline{b}(i) \end{bmatrix} + \underline{n}(i) \end{aligned} \qquad . \qquad (6.3.24)$$

If either $\underline{Q}(i) > \underline{0}$, $i = 0, 1, \ldots, k-1$, or $\underline{C}(i+1)\underline{R}(i)\underline{C}'(i+1) > \underline{0}$, $i = 0, 1, \ldots, k-1$ (or both), the unique $\underline{V}^*(i+i|i,U^*(0,i))$ which satisfies (6.3.21) to (6.3.23) is given by

$$\underline{V}^*(i+1|i,U^*(0,i)) = \tilde{\underline{\Delta}}(i|i,U^*(0,i))\tilde{\underline{C}}'(i+1)\cdot$$

$$[\underline{C}(i+1)\tilde{\underline{\Delta}}(i|i,U^*(0,k-1)\tilde{\underline{C}}'(i+1)+\underline{Q}(i+1)]^{-1}$$

$$i = 0, 1, \ldots, k-1 \qquad (6.3.25)$$

and the identification equations specify a Kalman filter for the augmented system $\tilde{\mathcal{S}}$.

In all cases, where the driving and/or observation noises may be degenerate, the conditional covariance, $\underline{\Sigma}(k|k,U^*(0,k-1))$, given by (6.3.21)-(6.3.23) is unique; and the conditional mean

$$\begin{bmatrix} \underline{\hat{x}}(k|k,U^*(0,k-1)) \\ \cdots \\ \underline{\hat{b}}(k|k,U^*(0,k-1)) \end{bmatrix}$$

is unique almost surely.(see chapter 3). Thus, we may assume that these quantities are known if $U^*(0,k-1)$ has been chosen. We can then formulate the following deterministic control problem at the kth-step:

Open-Loop Control Problem ($k \leq j \leq N-1$):

Given: $$\underline{\hat{x}}(j+1|k)^{\dagger} = \underline{A}(j)\underline{\hat{x}}(j|k) + \underline{b}(j|k)u(j) \tag{6.3.26}$$

$$\underline{\hat{b}}(j+1|k) = \underline{G}(j)\underline{\hat{b}}(j|k) \tag{6.3.27}$$

$$\underline{\Sigma}(j+1|k) = \underline{\tilde{A}}(j,u(j))\underline{\Sigma}(j|k)\underline{\tilde{A}}'(j,u(j)) + \underline{\tilde{R}}(j) \tag{6.3.28}$$

with known initial conditions at $j = k$

$$\underline{\hat{x}}(k|k) = \underline{\hat{x}}(k|k,U^*(0,k-1)) \quad ; \quad \underline{\hat{b}}(k|k) = \underline{\hat{b}}(k|k,U^*(0,k-1)) \quad ;$$

$$\underline{\Sigma}(k|k) = \underline{\Sigma}(k|k,U^*(0,k-1)) \quad . \tag{6.3.29}$$

We are to find a deterministic control sequence $U(k,N-1)$ such that it minimizes

$$J(U(k,N-1) \quad ; \quad U^*(0,k-1),k) = \frac{1}{2}\Big\{ \underline{\hat{x}}'(N|k)\underline{F}\,\underline{\hat{x}}(N|k) + tr\{\underline{\tilde{F}}\,\underline{\Sigma}(N|k)\} + \sum_{j=k}^{N-1} (\underline{\hat{x}}'(j|k)\underline{W}(j)\underline{\hat{x}}(j|k) + tr\{\underline{\tilde{W}}(j)\underline{\Sigma}(j|k)\} + h(j)u^2(j)\Big\} \tag{6.3.30}$$

subject to the constraints (6.3.26) to (6.3.28), where the matrices $\underline{\tilde{F}}$ and $\underline{\tilde{W}}(j)$ are defined by

---

† We shall not explicitly stress the dependence on the past control history $U^*(0,k-1)$; for this reason the symbol $U^*(0,k-1)$ shall be dropped without causing any confusion.

$$\tilde{F} = \begin{bmatrix} F & 0 \\ 0 & 0 \end{bmatrix} \quad ; \quad \tilde{W}(j) = \begin{bmatrix} W(j) & 0 \\ 0 & 0 \end{bmatrix} \tag{6.3.31}$$

For the above deterministic control problem, we shall denote its optimal control sequence by $U^o(k,N-1) \triangleq \{u^o(j|k)\}_{j=k}^{N-1}$, <u>where the superscript o is used to denote optimal for the open-loop control problem; the symbol $u^o(j|k)$ is used to indicate that the control is open-loop optimal conditioned on the observation up to the present time k.</u>

The solution for the above deterministic optimal control problem is given below; the detailed derivation will be carried out in section 6.4. The optimal control sequence, $U^o(k,N-1)$, is given by

$$u^o(j|k) = -[\tilde{h}(j|k)+\tilde{b}^{o\prime}(j|k)\tilde{K}(j+1|k)\tilde{b}^o(j|k)]^{-1}\cdot$$

$$\tilde{b}^{o\prime}(j|k)\tilde{K}(j+1|k)\,\Theta(j|k)\begin{bmatrix} x^o(j|k) \\ \cdots \\ z^o(j|k)\end{bmatrix} - \tilde{h}(j|k)d'(j+1)\begin{bmatrix} x^o(j|k) \\ \cdots \\ z^o(j|k)\end{bmatrix} \tag{6.3.32}$$

where $K(j|k)$, $j = k + 1, \ldots, N - 1$, satisfies the matrix difference equation

$$\tilde{K}(j|k) = \Theta'(j|k)\{\tilde{K}(j+1|k)-\tilde{K}(j+1|k)\tilde{b}^o(j|k)\{\tilde{h}(j|k)+\tilde{b}^{o\prime}(j|k)\tilde{K}(j+1|k)\tilde{b}^o(j|k)\}^{-1}\cdot$$

$$\tilde{b}^{o\prime}(j|k)\tilde{K}(j+1|k)\}\,\Theta(j|k)+\tilde{D}(j|k) \quad ; \quad K(N|k) = \begin{bmatrix} F & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \tag{6.3.33}$$

and for $j = k, \ldots, N - 1$

$$\underline{\mathcal{Q}}(j|k) = \hat{\underline{A}}(j) - \tilde{\underline{b}}^o(j|k)\tilde{h}^{-1}(j|k)\underline{d}'(j+1) \quad ;$$

$$\hat{\underline{A}}(j) = \begin{bmatrix} \underline{A}(j) & \underline{0} & \cdots & \underline{0} \\ \underline{0} & \underline{A}(j)g_{11} & \cdots & \underline{A}(j)g_{n1} \\ \cdots & \cdots & \cdots & \cdots \\ \underline{0} & \underline{A}(j)g_{1n} & \cdots & \underline{A}(j)g_{nn} \end{bmatrix} \tag{6.3.34}$$

$$\tilde{\underline{D}}(j|k) = \tilde{\underline{W}}(j) - \underline{d}(j+1)\tilde{h}^{-1}(j|k)\underline{d}'(j+1) \tag{6.3.35}$$

$$\underline{d}(j) = \begin{bmatrix} \underline{0} \\ \underline{A}(j-1)\underline{S}(j)\underline{e}_1 \\ \vdots \\ \underline{A}(j-1)\underline{S}(j)\underline{e}_n \end{bmatrix} \varepsilon\ R^{n(n+1)} \quad ; \quad \tilde{\underline{b}}^o(j|k) = \begin{bmatrix} \hat{\underline{b}}^o(j|k) \\ \underline{\Sigma}_b^o(j|k)\underline{G}'(j)\underline{e}_1 \\ \vdots \\ \underline{\Sigma}_b^o(j|k)\underline{G}'(j)\underline{e}_n \end{bmatrix} \varepsilon\ R^{n(n+1)} \quad ;$$

$$\underline{c}^o(j|k) = \begin{bmatrix} \underline{\Sigma}_{xb}^o(j|k)\underline{e}_1 \\ \vdots \\ \underline{\Sigma}_{xb}^o(j|k)\underline{e}_n \end{bmatrix} \tag{6.3.36}$$

$$\tilde{h}(j|k) = h(j) + tr\{\underline{\Sigma}_b^o(j|k)\underline{S}(j+1)\} \quad . \tag{6.3.37}$$

The matrices $\underline{\Sigma}_b^o(j|k)$, $\underline{S}(j + 1)$, $j = k, k + 1, \ldots, N - 1$, are given by

$$\begin{aligned} \underline{\Sigma}_b^o(j+1|k) &= \underline{G}(j)\underline{\Sigma}_b^o(j|k)\underline{G}'(j)+\underline{N}(j) \qquad j = k, \ldots, N - 1 \quad ; \\ \underline{\Sigma}_b^o(k|k) &= \underline{\Sigma}_b(k|k,\underline{U}^*(0,k-1)) \end{aligned} \tag{6.3.38}$$

$$\underline{S}(j) = \underline{A}'(j)\underline{S}(j+1)\underline{A}(j) + \underline{W}(j) \qquad j = k, \ldots, N - 1 \quad ; \quad \underline{S}(N) = \underline{F} \tag{6.3.39}$$

and the vector $\hat{\underline{b}}^o(j|k)$ satisfies

---

†The vectors $\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n$ represent the natural basis in R_n.

$$\underline{b}^o(j+1|k) = \underline{G}(j)\underline{b}^o(j|k) \quad , \quad j = k, \ldots, N-1 \quad ; \quad \underline{b}^o(k|k) = \hat{\underline{b}}(k|k,U^*(0,k-1)) \; . \tag{6.3.40}$$

To find the O.L.F.O. control sequence, we have to solve the above open loop control problem for $k = 0, 1, \ldots$. The O.L.F.O. control $\{u^*(k)\}_{k=0}^{N-1}$ is then given by

$$u^*(k) = u^o(k|k) \qquad k = 0, 1, \ldots, N-1 \tag{6.3.41}$$

where $u^o(k|k)$ is given by (6.3.32) to (6.3.40). The structure of the O.L.F.O. control system is described by Figure 6.1. Though the equations are complicated, the digital computer implementation of O.L F.O. control sequence is actually straightforward. A flow chart description of the O.L.F.O. control is given in Figures 6.2 and 6.3. In the following, we shall outline the computational procedure to find the O.L.F.O. control sequence.

1. If $k = 0$, $\underline{y}(0)$ is observed, and $\hat{\underline{x}}(0|0,U^*(0,-1))$, $\hat{\underline{b}}(0|0,U^*(0,-1))$, $\underline{\Sigma}(0|0,U^*(0,-1))$ are given by (6.3.19) and (6.3.23). If $k > 0$, assume that $U^*(0,k-1)$ is chosen and $Y_{U^*(0,k-1)}(0,k)$ is observed; compute $\hat{\underline{x}}(k|k,U^*(0,k-1))$, $\hat{\underline{b}}(k|k,U^*(0,k-1))$, and $\underline{\Sigma}(k|k,U^*(0,k-1))$ using the identification equations (6.3.19) to (6.3.23).
2. Compute $\underline{\Theta}(j|k)$, $\tilde{\underline{D}}(j|k)$, $\tilde{\underline{b}}^o(j|k)$, $\tilde{h}(j|k)$ for $j = k, k+1, \ldots, N-1$ using equations (6.3.34) to (6.3.40).
3. Compute $\underline{K}(k+1|k)$ using (6.3.33), and the O.L.F.O. control to be applied at step k is given by (6.3.41).

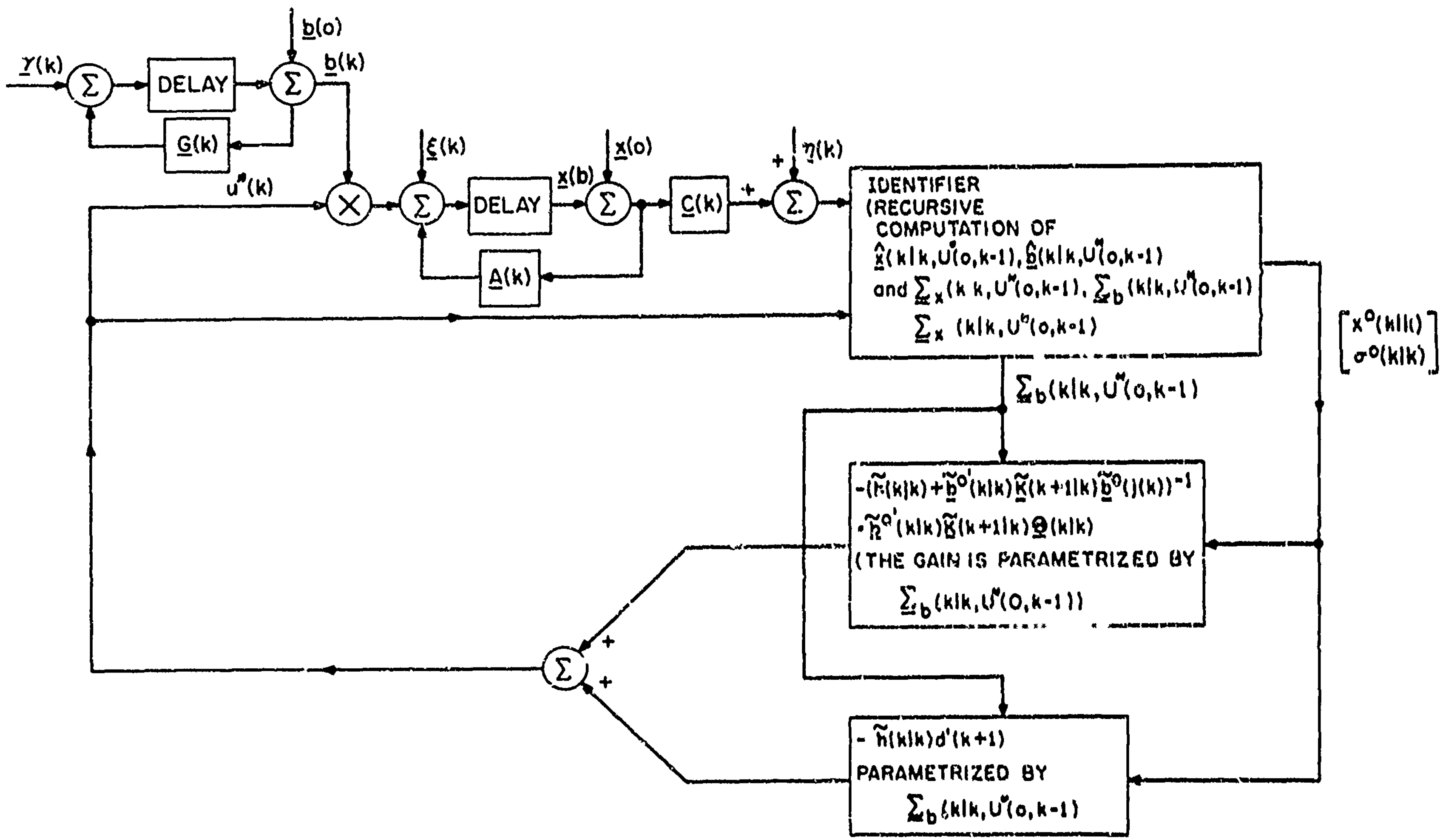


Fig. 6.1 THE STRUCTURE OF O.L.F.O. CONTROL SYSTEM. THE OBSERVATION NOISE MAY ASSUME TO BE NONDEGENERATE, DEGENERATE OR TOTALLY SINGULAR.

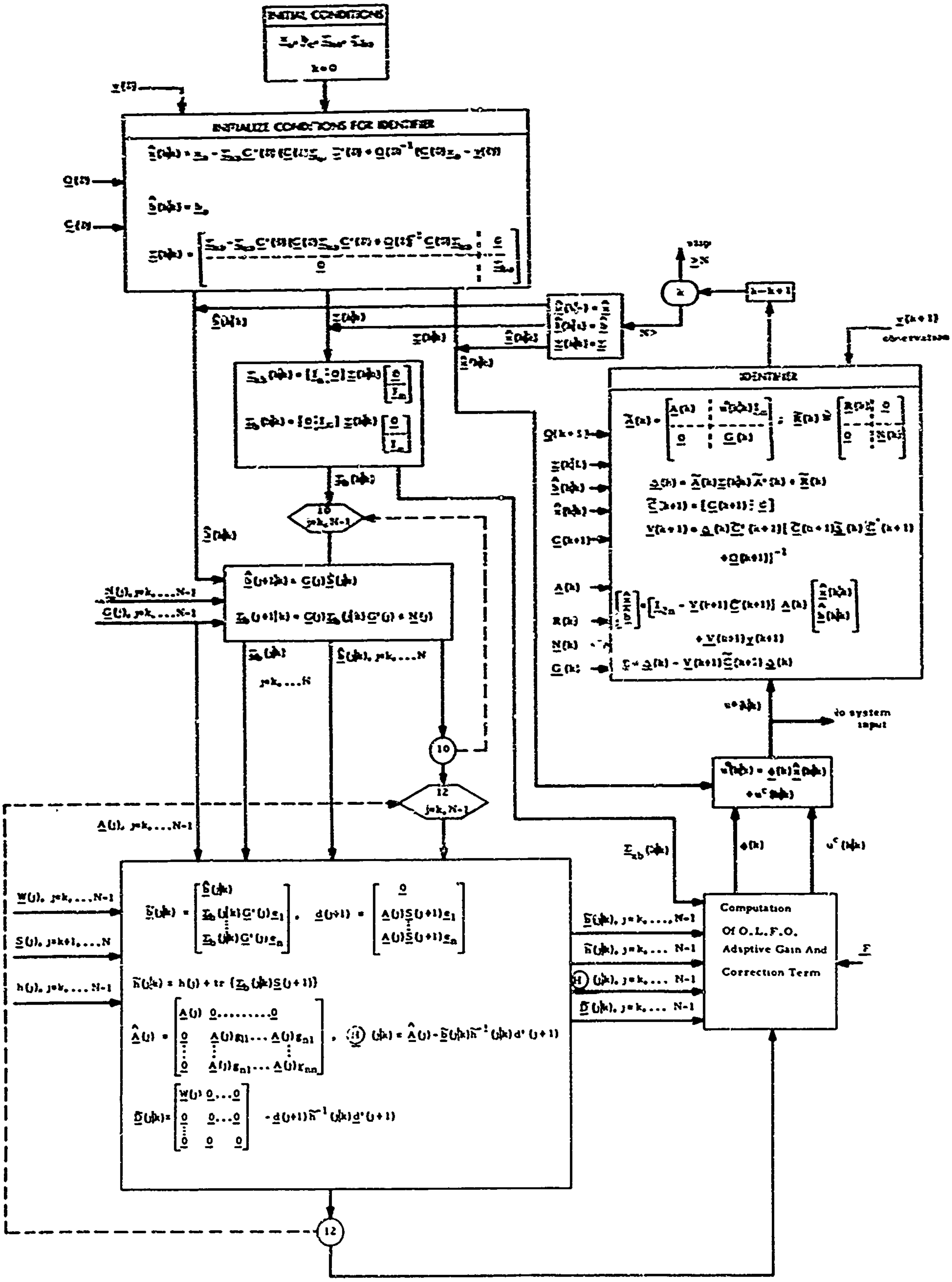


Fig. 6.2 FLOW CHART FOR ON-LINE COMPUTATION OF O.L.F.O. CONTROL

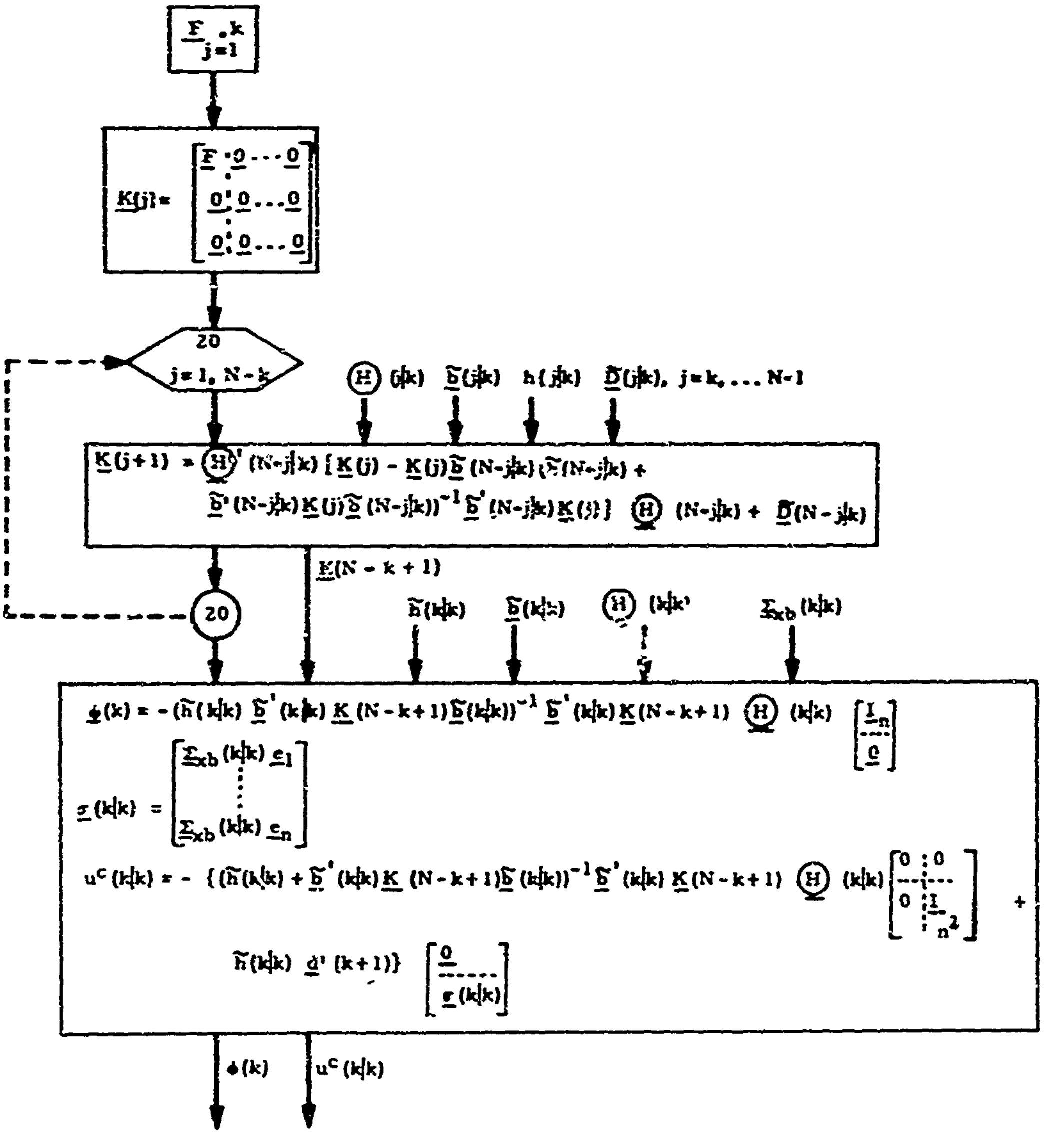


Fig. 6.3 FLOW CHART FOR COMPUTING THE O.L.F.O. ADAPTIVE GAIN AND CORRECTION TERM

4. Advance $k \to k + 1$ and repeat 1 through 3 until $k = N - 1$.

We note that the O.L.F.O. control sequence $U^*(0,N - 1)$ is adaptive in nature.

Before we go into the derivation of the O.L.F.O. control, let us first look into the solution carefully and discuss some of its implications.

In essence, we are forcing some sort of "separation" in our formulation. The overall control problem is split into an identification and a deterministic control problem. However, the effect of the identification error will be taken into account in the deterministic control problem. Thus, this does not correspond to pure separation as it is in the case of stochastic control of linear system with known dynamics(chapter 5).

Let us first look into the identification equations (6.3.19)-(6.3.23). Suppose that $\underline{Q}(k) > \underline{0}$. If $u(i) = 0$, then from (6.3.21), we have

$$\underline{V}^*(i+1|i,U(0,i)) = \begin{bmatrix} \underline{V}_1^*(i+1|i,U(0,i)) \\ \cdots \\ \underline{0} \end{bmatrix} \tag{6.3.42}$$

and so (6.3.19) implies

$$\hat{\underline{b}}(i+1|i+1,U(0,i)) = \underline{G}(i)\hat{\underline{b}}(i|i,U(0,i-1)) \quad . \tag{6.3.43}$$

Therefore, a nonzero input is necessary to identify the gain parameter vector $\underline{b}(k)$. From the equation (6.2.1), we see that if $u(i)$ is very large, then for the most part the value of $\underline{x}(i + 1)$ will be due to $\underline{b}(k)u(k)$, and so the observation $y(k)$ will contain a large amount of information about the gain parameter $\underline{b}(k)$. Therefore, we would expect that large input

magnitudes will be helpful for the identification of $\underline{b}(k)$. For a control sequence $\{u(k)\}_{k=0}^{N-1}$, if its total energy is high, we would expect such control sequence to be useful for identification purpose. But large control energy will also give rise to a high cost (6.3.1). From the control point of view, we would like to use just enough control energy to regulate the state of the system. In general, there is a conflict between identification and control, and <u>a reasonable control sequence should appropriately distribute its total energy to identify and/or control of the system $\mathcal{S}$</u>.

Let us consider (6.3.32)-(6.3.33). Comparing with the Levis'[75] results, we note that $u^o(j|k)$ is the optimal control for the problem of controlling the system $\tilde{S}_k$:

$$\tilde{S}_k: \quad \underline{\tilde{x}}(j+1|k) = \underline{\hat{A}}(j)\underline{\tilde{x}}(j|k) + \underline{\tilde{b}}(j|k)u(j|k) \quad ; \quad \underline{\tilde{x}}(j+1|k) \triangleq \begin{bmatrix} \underline{\hat{x}}^o(j+1|k) \\ \cdots \\ \underline{c}^o(j+1|k) \end{bmatrix} \tag{6.3.44}$$

with the cost criteria

$$J = \underline{\tilde{x}}'(N|k)\underline{\tilde{F}}\ \underline{\tilde{x}}(N|k)+\sum_{i=k}^{N-1}\{\underline{\tilde{x}}(i|k)\underline{\tilde{W}}(i|k)\underline{\tilde{x}}(i|k)+\tilde{h}(j|k)u^2(j|k)+2\underline{\tilde{x}}'(i|k)\underline{d}(i+1)u(i|k)\} \tag{6.3.45}$$

Therefore we can visualize $\tilde{h}(j|k)$ as the <u>modified relative weighting on the control</u>. From (6.3.31), we note that $h(j|k)$ relates in a direct manner with $\underline{\Sigma}_b^o(j|k)$. In a statistical sense, $\underline{\Sigma}_b(k|k)$ reflects the level of confidence we have about the estimate of $\underline{b}(k)$. <u>The modification on the relative weighting on the control is such that heavy weighting is put in the control if we have little confidence on the estimate</u> of $\underline{b}(k)$; therefore,

the control action will be very cautious and control energy will not be used unless it is very necessary.

Let us write

$$u^*(k|k) = -(\tilde{h}(k|k)+\underline{\tilde{b}}^{o\prime}(k|k)\underline{\tilde{K}}(k+1|k)\underline{\tilde{b}}^{o}(k|k))^{-1}\cdot\underline{\tilde{b}}^{o\prime}(k|k)\underline{\tilde{K}}(k+1|k)\,\underline{\textcircled{H}}(k|k)$$

$$\begin{bmatrix} \underline{I}_n & \vdots & \underline{0} \\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{0} \end{bmatrix}\begin{bmatrix} \underline{\hat{x}}^o(k|k) \\ \cdots \\ \underline{\sigma}^o(k|k) \end{bmatrix} - (\tilde{h}(k|k)+\underline{\tilde{b}}^{o\prime}(k|k)\underline{\tilde{K}}(k+1|k)\underline{\tilde{b}}^{o\prime}(k|k))^{-1}\underline{\tilde{K}}(k+1|k)$$

$$\underline{\textcircled{H}}(k|k)\begin{bmatrix} \underline{0} & \vdots & \underline{0} \\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{I}_n \end{bmatrix} + \tilde{h}^{-1}(k|k)\underline{d}'(k+1)\begin{bmatrix} \underline{\hat{x}}^o(k|k) \\ \cdots \\ \underline{\sigma}^o(k|k) \end{bmatrix} \quad . \tag{6.3.46}$$

We shall call the row vector (1xn)

$$\underline{\gamma}(k) \triangleq -(\tilde{h}(k|k)+\underline{\tilde{b}}^{o\prime}(k|k)\underline{\tilde{K}}(k+1|k)\underline{\tilde{b}}^{o}(k|k))^{-1}\cdot\underline{\tilde{b}}^{o\prime}(k|k)\underline{\tilde{K}}(k+1|k)\,\underline{\textcircled{H}}(k|k)\begin{bmatrix} \underline{I}_n \\ \cdots \\ \underline{0} \end{bmatrix} \tag{6.3.47}$$

the <u>O.L.F.O. adaptive gain</u>, and the term

$$u^c(k|k) = -\{(\tilde{h}(k|k)+\underline{\tilde{b}}^{o\prime}(k|k)\underline{\tilde{K}}(k+1|k)\underline{\tilde{b}}^{o}(k|k)\}^{-1}\cdot\underline{\tilde{b}}^{o\prime}(k|k)\underline{\tilde{K}}(k+1|k)\,\underline{\textcircled{H}}(k|k)$$

$$\begin{bmatrix} \underline{0} & \vdots & \underline{0} \\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{I}_n \end{bmatrix} + \tilde{h}(k|k)\underline{d}'(k+1)\begin{bmatrix} \underline{\hat{x}}^o(k|k) \\ \cdots \\ \underline{\sigma}^o(k|k) \end{bmatrix} \tag{6.3.48}$$

the <u>correction term</u>. Thus, the <u>O.L.F.O.</u> control, (6.3.46), becomes

$$u^*(k|k) = \underline{\gamma}(k|k)\underline{\hat{x}}^o(k|k) + u^c(k|k) \quad . \tag{6.3.49}$$

From (6.3.33)-(6.3.37), we note that $\underline{\Sigma}_b(k|k,U^*(0,k-1))$ affects indirectly the O.L.F.O. adaptive gain and the correction term. The cross-error

covariance, $\underline{\Sigma}_{xb}(k|k,U^*(0,k-1))$, only affects the correction term; and if $\underline{\Sigma}_{xb}(k|k,U^*(0,k-1))$ is zero, then from (6.3.36) and (6.3.48) we conclude $u^c(k|k) = 0$.

Assume that $\underline{\Sigma}_b(k|k,U^*(0,k-1)) = \underline{0}$, then from (6.3.33)-(6.3.37), we have inductively

$$\tilde{\underline{K}}(k|k) = \begin{bmatrix} \underline{K}(k) & \vdots & \underline{0} \\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{0} \end{bmatrix} \tag{6.3.50}$$

where $\underline{K}(k)$ is given by (5.3.6), and from (6.3.47), the O.L.F.O. adaptive gain is

$$\underline{c}(k) = -(h(k) + \underline{b}'(k)\underline{K}(k+1)\underline{b}(k))^{-1}\underline{b}'(k)\underline{K}(k+1)\underline{A}(k) \tag{6.3.51}$$

which is the <u>truly optimum gain</u> (see chapter 5, section 5.3). The assumption that $\underline{\Sigma}_b(k|k,U^*(0,k-1)) = \underline{0}$ also implies $\underline{\Sigma}_{xb}(k|k,U^*(0,k-1)) = \underline{0}$, and so the correction term is zero, and

$$u^*(k|k) = -(h(k) + \underline{b}'(k)\underline{K}(k+1)\underline{b}(k))^{-1}\underline{b}'(k)\underline{K}(k+1)\underline{A}(k)\hat{\underline{x}}^o(k|k) \quad . \tag{6.3.52}$$

Thus we see that <u>if for some k, the identification of b(k) is assured to be exact</u>, i.e., <u>the level of confidence on the estimated gain parameters is very very high, then the O.L.F.O. control will act optimally and use the obtained estimate of b(k) as if it were the true gain vector.</u>

Finally, we would like to comment on the computational requirements of the proposed scheme. The computation of the O.L.F.O. control is done on-line. At each time unit k, we have to solve a one step 2n-vector difference equation and a one step $2n \times 2n$ matrix difference equation,

(6.3.19)-(6.3.23); computing the parameters, (6.3.34)-(6.3.40), which involve some one step computation, (6.3.34)-(6.3.37), and an $N-k$ steps n-vector difference equation (6.3.40) and an $N-k$ steps $n\times m$ matrix difference equation (6.3.38); finally we have to solve an $N-k$ step $(m+1)n \times (n+1)n$ matrix difference equation (6.3.33). (Note that the matrix difference equation (6.3.39) can be precomputed off-line.) The O.L.F.O. control is then computed using (6.3.32). The total storage capacity needed corresponds to the storage of the state and parameter estimates (2n) and the error covariance matrix ($2n \times 2n$). The capability of computing the O.L.F.O. control sequence in almost real time will depend on the complexity of the system being considered and the computation speed of the digital computer used to implement the O.L.F.O. control (see also chapter 7).

6.4 Open-Loop Optimal Control

In this section, we shall derive the open-loop optimal control for the deterministic control problem (6.3.26)-(6.3.31). The deterministic formulation allows us to use the discrete matrix minimum principle (theorem 2.4.1) to derive the set of necessary conditions for optimality.

Let us form the Hamiltonian for the deterministic control problem (6.3.26)-(6.3.31) for $j = k, k+1, \ldots, N-1$.

$$
\begin{aligned}
H_{j|k} = {} & <\underline{p}_x(j+1|k), \underline{A}(j)\hat{\underline{x}}(j|k)+\hat{\underline{b}}(j|k)u(j)-\hat{\underline{x}}(j|k)> + <\underline{p}_b(j+1|k), \underline{G}(j)\hat{\underline{b}}(j|k)- \\
& \hat{\underline{b}}(j|k)> + \mathrm{tr}\left\{(\tilde{\underline{A}}(j,u(j))\underline{\Sigma}(j|k)\tilde{\underline{A}}'(j,u(j))+\tilde{\underline{R}}(j)-\underline{\Sigma}(j|k))\underline{P}'(j+1|k)\right\} \\
& + \frac{1}{2}<\hat{\underline{x}}(j|k), \underline{W}(j)\hat{\underline{x}}(j|k)> + \frac{1}{2}h(j)u^2(j) + \frac{1}{2}\,\mathrm{tr}\,\tilde{\underline{W}}(j)\underline{\Sigma}(j|k) \qquad (6.4.1)
\end{aligned}
$$

where $\underline{p}_x(j|k)$ is the costate vector associated with $\hat{\underline{x}}(j|k)$, $\underline{p}_b(j|k)$ is the costate vector associated with $\hat{\underline{b}}(j|k)$ and $\underline{P}(j|k)$ is the costate matrix associated with $\underline{\Sigma}(j|k)$. Use of the discrete matrix minimum principle leads to the following relations

(a) The canonical equations are:

$$\hat{\underline{x}}^o(j+1|k) = \underline{A}(j)\hat{\underline{x}}^o(j|k) + \hat{\underline{b}}^o(j|k)u^o(j|k) \tag{6.4.2}$$

$$\hat{\underline{b}}^o(j+1|k) = \underline{G}(j)\hat{\underline{b}}^o(j|k) \tag{6.4.3}$$

$$\underline{\Sigma}^o(j+1|k) = \tilde{\underline{A}}(j,u^o(j|k))\underline{\Sigma}^o(j|k)\tilde{\underline{A}}'(j,u^o(j|k)) + \tilde{\underline{R}}(j) \tag{6.4.4}$$

$$\underline{p}_x^o(j|k) = \underline{A}'(j)\underline{p}_x^o(j+1|k) + \underline{W}(j)\underline{x}^o(j|k) \tag{6.4.5}$$

$$\underline{p}_b^o(j|k) = \underline{G}'(j)\underline{p}_b^o(j+1|k) + \underline{p}_x^o(j+1|k)u^o(j|k) \tag{6.4.6}$$

$$\underline{P}^o(j|k) = \tilde{\underline{A}}'(j,u^o(j|k))\underline{P}^o(j+1|k)\tilde{\underline{A}}(j,u^o(j|k)) + \frac{1}{2}\,\underline{W}(j) \tag{6.4.7}$$

(b) The boundary conditions are:

at time k: $\hat{\underline{x}}^o(k|k) = \hat{\underline{x}}(k|k,U^*(0,k-1));\hat{\underline{b}}^o(k|k) = \hat{\underline{b}}(k|k,U^*(0,k-1))$;

$$\underline{\Sigma}^o(k|k) = \underline{\Sigma}(k|k,U^*(0,k-1)) \tag{6.4.8}$$

at time N: $\underline{p}_x^o(N|k) = F\,\hat{x}^o(N|k);\underline{p}_b^o(N|k) = \underline{0}$;

$$\underline{P}^o(N|k) = \frac{1}{2}\,\tilde{\underline{F}} \tag{6.4.9}$$

(c) In minimizing the Hamiltonian, we set (for $j \geq k$)

$$\left.\frac{\partial H}{\partial u(j)}\right|_{u^o(j|k)} = \hat{\underline{b}}^{o\prime}(j|k)\underline{p}_x^o(j+1|k)+2\mathrm{tr}\{\underline{P}_{11}^o(j+1|k)\underline{A}(j)\underline{\Sigma}_{xb}^o(j|k)+$$

$$\underline{G}(j)\underline{\Sigma}_b^o(j|k)\underline{P}_{12}^o(j+1|k)\}+u^o(j|k)\cdot\{h(j)+\mathrm{tr}[2\underline{\Sigma}_b^o(j|k)\underline{P}_{11}^o(j+1|k)]\} = 0$$

(6.4.10)

where we have decomposed the costate matrix as follows

$$\underline{P}^{o}(j|k) = \begin{bmatrix} \underline{P}^{o}_{11}(j|k) & \vdots & \underline{P}^{o}_{12}(j|k) \\ \cdots & & \cdots \\ \underline{P}^{o}_{21}(j|k) & \vdots & \underline{P}^{o}_{22}(j|k) \end{bmatrix} \tag{6.4.11}$$

From (6.4.7) and (6.4.9), we deduced that $\underline{P}^{o}(j|k)$ is nonnegative definite since $\underline{F}$ and $\underline{W}(j)$ are assumed to be nonnegative definite. Therefore,

$$\left.\frac{\partial^2 H}{\partial u^2(j)}\right|_{u^{o}(j|k)} = h(j) + 2\ \mathrm{tr}\{\underline{\Sigma}^{o}_{b}(j|k)\underline{P}^{o}_{11}(j+1|k)\} > 0 \tag{6.4.12}$$

and so the control $u^{o}(j|k)$ given by

$$u^{o}(j|k) = -[h(j)+2\mathrm{tr}(\underline{\Sigma}^{o}_{b}(j|k)\underline{P}^{o}_{11}(j+1|k))]^{-1}\{\hat{\underline{b}}^{o\prime}(j|k)\underline{p}^{o}_{x}(j+1|k)+$$

$$2\mathrm{tr}(\underline{P}^{o}_{11}(j+1|k)\underline{A}(j)\underline{\Sigma}^{o}_{xb}(j|k)+\underline{G}(j)\underline{\Sigma}^{o}_{b}(j|k)\underline{P}^{o}_{12}(j+1|k))\}\ ;$$

$$j \geq k \tag{6.4.13}$$

indeed minimizes the Hamiltonian.

From (6.4.4) we obtain equations for $\underline{\Sigma}^{o}_{xb}(j|k)$ and $\underline{\Sigma}^{o}_{b}(j|k)$ for $j \geq k$:

$$\underline{\Sigma}^{o}_{xb}(j+1|k) = \underline{A}(j)\underline{\Sigma}^{o}_{xb}(j|k)\underline{G}'(j) + u^{o}(j|k)\underline{\Sigma}^{o}_{b}(j|k)\underline{G}'(j) \tag{6.4.14}$$

$$\underline{\Sigma}^{o}_{b}(j+1|k) = \underline{G}(j)\underline{\Sigma}^{o}_{b}(j|k)\underline{G}'(j) + \underline{N}(j) \tag{6.4.15}$$

with initial conditions

$$\underline{\Sigma}^{o}_{xb}(k|k) = \underline{\Sigma}_{xb}(k|k,U^{*}(0,k-1))\quad ;\quad \underline{\Sigma}^{o}_{b}(k|k) = \underline{\Sigma}_{b}(k|k,U^{*}(0,k-1))\ . \tag{6.4.16}$$

From equations (6.4.7) and (6.4.9), we obtain the equations for $\underline{P}_{11}^{o}(j|k)$ and $\underline{P}_{12}^{o}(j|k)$ for $j = k, k+1, \ldots, N$:

$$\underline{P}_{11}^{o}(j|k) = \underline{A}'(j)\underline{P}_{11}^{o}(j+1|k)\underline{A}(j) + \frac{1}{2}\underline{W}(j) \tag{6.4.17}$$

$$\underline{P}_{12}^{o}(j|k) = \underline{A}'(j)\underline{P}_{12}^{o}(j+1|k)\underline{G}(j) + u^{o}(j|k)\underline{A}'(j)\underline{P}_{11}^{o}(j+1|k) \tag{6.4.18}$$

$$\underline{P}_{11}^{o}(N|k) = \frac{1}{2}\underline{F} \quad ; \quad \underline{P}_{12}^{o}(N|k) = \underline{0} \quad . \tag{6.4.19}$$

We note from (6.4.13) to (6.4.19), that the values of $\underline{p}_{b}^{o}(j|k)$ and $\underline{P}_{22}^{o}(j|k)$ are irrelevant in computing the open-loop optimal control sequence $u^{o}(j|k)$. From (6.4.17), $\underline{P}_{11}^{o}(j|k)$ is independent of the observation and the control, and thus it can be precomputed. To emphasize this fact, we define

$$\underline{S}(j) \triangleq 2\underline{P}_{11}^{o}(j|k) \geq \underline{0} \quad ; \quad N \geq j \geq 0 \tag{6.4.20}$$

and so $\underline{S}(j)$ satisfies the matrix difference equation

$$\underline{S}(j) = \underline{A}'(j)\underline{S}(j+1)\underline{A}(j) + \underline{W}(j) \quad ; \quad \underline{S}(N) = \underline{F} \quad . \tag{6.4.21}$$

From (6.4.15), $\underline{\Sigma}_{b}^{o}(j|k)$ only depends on the observation; thus it is meaningful to define a "modified control weighting":

$$\tilde{h}(j|k) \triangleq h(j) + \mathrm{tr}(\underline{\Sigma}_{b}^{o}(j|k)\underline{S}(j+1)) > 0 \quad . \tag{6.4.22}$$

Let us define

$$\underline{\phi}^{o}(j|k) = \begin{bmatrix} 2\underline{P}_{12}^{o}(j|k)\underline{e}_{1} \\ \vdots \\ 2\underline{P}_{12}^{o}(j|k)\underline{e}_{n} \end{bmatrix} \varepsilon\ R^{n^2} \quad . \tag{6.4.23}$$

Then by using (6.3.34) to (6.3.36), (6.4.2), (6.3.13) to (6.4.23), we obtain the set of matrix difference equations: $j = k, k+1, \ldots, N-1$.

$$\begin{bmatrix} \underline{\hat{x}}^{o}(j+1|k) \\ \cdots \\ \underline{\sigma}^{o}(j+1|k) \end{bmatrix} = \underline{\text{Ⓗ}}(j|k) \begin{bmatrix} \underline{\hat{x}}^{o}(j|k) \\ \cdots \\ \underline{\sigma}^{o}(j|k) \end{bmatrix} - \underline{\tilde{b}}^{o}(j|k)\tilde{h}^{-1}(j|k)\underline{\tilde{b}}^{o\prime}(j|k) \begin{bmatrix} \underline{p}_{x}^{o}(j+1|k) \\ \cdots \\ \underline{\phi}^{o}(j+1|k) \end{bmatrix} \tag{6.4.24}$$

$$\begin{bmatrix} \underline{p}_{x}^{o}(j|k) \\ \cdots \\ \underline{\phi}^{o}(j|k) \end{bmatrix} = \underline{\text{Ⓗ}}'(j|k) \begin{bmatrix} \underline{p}_{x}^{o}(j+1|k) \\ \cdots \\ \underline{\phi}^{o}(j+1|k) \end{bmatrix} + \underline{\tilde{D}}(j|k) \begin{bmatrix} \underline{\hat{x}}^{o}(j|k) \\ \cdots \\ \underline{\sigma}^{o}(j|k) \end{bmatrix} \tag{6.4.25}$$

with boundary conditions at time N:

$$\begin{bmatrix} \underline{p}_{x}^{o}(N|k) \\ \cdots \\ \underline{\phi}^{o}(N|k) \end{bmatrix} = \begin{bmatrix} \underline{F} & \underline{0} & \cdots & \underline{0} \\ \underline{0} & \underline{0} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \underline{0} & \underline{0} & \cdots & \underline{0} \end{bmatrix} \begin{bmatrix} \underline{\hat{x}}^{o}(N|k) \\ \cdots \\ \underline{\sigma}^{o}(N|k) \end{bmatrix} . \tag{6.4.26}$$

From (6.4.24) to (6.4.26) and (6.4.3), (6.4.16), we can solve for $\underline{p}_{x}^{o}(j|k)$, $\underline{\phi}^{o}(j|k)$, $\underline{\hat{x}}^{o}(j|k)$ and $\underline{\sigma}^{o}(j+1|k)$. To "bypass" the two point boundary value problem, we define the matrix $\underline{\tilde{K}}(j|k)$ by

$$\begin{bmatrix} \underline{p}_{x}^{o}(j|k) \\ \\ \underline{\phi}^{o}(j|k) \end{bmatrix} = \underline{\tilde{K}}(j|k) \begin{bmatrix} \underline{\hat{x}}^{o}(j|k) \\ \\ \underline{\sigma}^{o}(j|k) \end{bmatrix} . \tag{6.4.27}$$

Substituting (6.4.27) into (6.4.24) and (6.4.25) we obtain

$$[\underline{I} + \underline{\tilde{b}}^{o}(j|k)\tilde{h}^{-1}(j|k)\underline{\tilde{b}}^{o\prime}(j|k)\underline{\tilde{K}}(j+1|k)] \begin{bmatrix} \underline{\hat{x}}^{o}(j+1|k) \\ \\ \underline{\sigma}^{o}(j+1|k) \end{bmatrix} = \underline{\text{Ⓗ}}(j|k) \begin{bmatrix} \underline{\hat{x}}^{o}(j|k) \\ \\ \underline{\sigma}^{o}(j|k) \end{bmatrix} \tag{6.4.28}$$

$$[\underline{\tilde{K}}(j|k) - \underline{\tilde{D}}(j|k)] \begin{bmatrix} \underline{\hat{x}}^{o}(j|k) \\ \cdots \\ \underline{\sigma}^{o}(j|k) \end{bmatrix} = \underline{\text{Ⓗ}}'(j|k)\underline{\tilde{K}}(j+1|k) \begin{bmatrix} \underline{\hat{x}}^{o}(j+1|k) \\ \cdots \\ \underline{\sigma}^{o}(j+1|k) \end{bmatrix} . \tag{6.4.29}$$

If $[\underline{I} + \tilde{\underline{b}}^{o}(j|k)\tilde{h}^{-1}(j|k)\tilde{\underline{b}}^{o\prime}(j|k)\tilde{\underline{K}}(j+1|k)]$ has an inverse, then (6.4.28), (6.4.29) imply that

$$\{\tilde{\underline{K}}(j|k)-\tilde{\underline{D}}(j|k)-\underline{\Theta}'(j|k)\tilde{\underline{K}}(j+1|k)[\underline{I}+\tilde{\underline{b}}^{o}(j|k)\tilde{h}^{-1}(j|k)\tilde{\underline{b}}^{o\prime}(j|k)\tilde{\underline{K}}(j+1|k)]^{-1}\underline{\Theta}(j|k)\}\cdot$$

$$\begin{bmatrix} \hat{\underline{x}}^{o}(j|k) \\ \\ \underline{\sigma}^{o}(j|k) \end{bmatrix} = \underline{0} \quad . \tag{6.4.30}$$

Since $\hat{\underline{x}}^{o}(j|k)$, $\underline{\sigma}^{o}(j|k)$ can be arbitrary, (6.4.30) implies that the matrix difference equation holds:

$$\tilde{\underline{K}}(j|k) = \underline{\Theta}'(j|k)(\tilde{\underline{K}}(j+1|k)-\tilde{\underline{K}}(j+1|k)\tilde{\underline{b}}^{o}(j|k)\{\tilde{h}(j|k)+\tilde{\underline{b}}^{o\prime}(j|k)\tilde{\underline{K}}(j+1|k)\tilde{\underline{b}}^{o}(j|k)\}^{-1}\cdot$$

$$\tilde{\underline{b}}^{o\prime}(j|k)\tilde{K}(j+1|k))\underline{\Theta}(j|k)+\tilde{\underline{D}}(j|k) \quad ; \quad \tilde{\underline{K}}(N|k) = \begin{bmatrix} \underline{F} & \underline{0} & \cdots & \underline{0} \\ \underline{0} & \underline{0} & \cdots & \underline{0} \\ \cdots & \cdots & \cdots & \cdots \\ \underline{0} & \underline{0} & \cdots & \underline{0} \end{bmatrix}$$

(6.4.31)

where we have used the matrix identity:[66]

$$(\underline{I}_n + \underline{A}\,\underline{B}')^{-1} = \underline{I}_n - \underline{A}(\underline{I}_r + \underline{B}'\underline{A})^{-1}\underline{B}' \quad ; \quad \underline{A},\underline{B} \in M_{nr} \quad . \tag{6.4.32}$$

The identity (6.4.32) is true provided one of the inverses exists. The two point boundary problem is now transformed to the problem of finding a solution of the matrix difference equation (6.4.31). The existence and uniqueness of (6.4.24) to (6.4.26) and (6.4.8), (6.4.16) can be deduced from the existence and uniqueness of $\tilde{\underline{K}}(j|k)$, $N \geq j \geq k$, satisfying (6.4.31).

The optimal open-loop control is given by: $(N - 1 \geq j \geq k)$

$$u^o(j|k) = -\tilde{h}^{-1}(j|k)\ \tilde{\underline{b}}^{o\prime}(j|k)\begin{bmatrix}\underline{p}_x^o(j+1|k)\\ \cdots \\ \underline{\delta}^o(j+1|k)\end{bmatrix} + \underline{d}'(j+1)\begin{bmatrix}\hat{\underline{x}}^o(j|k)\\ \cdots \\ \underline{\sigma}^o(j|k)\end{bmatrix}$$

$$= -\tilde{h}^{-1}(j|k)\tilde{\underline{b}}^{o\prime}(j|k)\tilde{\underline{K}}(j+1|k)[\underline{I}+\tilde{\underline{b}}^o(j|k)\tilde{h}^{-1}(j|k)\tilde{\underline{b}}^{o\prime}(j|k)\tilde{\underline{K}}(j+1|k)]^{-1}\cdot$$

$$\underline{\text{Ⓗ}}(j|k)\begin{bmatrix}\hat{\underline{x}}^o(j|k)\\ \cdots \\ \underline{\sigma}^o(j|k)\end{bmatrix} - \tilde{h}^{-1}(j|k)\underline{d}'(j+1)\begin{bmatrix}\hat{\underline{x}}^o(j|k)\\ \cdots \\ \underline{\sigma}^o(j|k)\end{bmatrix} \qquad (6.4.33)$$

Using (6.4.32) and also the matrix identity[66]

$$\underline{I} - \underline{A}(\underline{A} + \underline{B})^{-1} = \underline{B}(\underline{B} + \underline{A})^{-1} \qquad (6.4.34)$$

(6.4.33) becomes $(N - 1 \geq j \geq k)$

$$u^o(j|k) = -(\tilde{h}(j|k)+\tilde{\underline{b}}^{o\prime}(j|k)\tilde{\underline{K}}(j+1|k)\tilde{\underline{b}}^o(j|k))^{-1}\tilde{\underline{b}}^{o\prime}(j|k)\tilde{\underline{K}}(j+1|k)\ \underline{\text{Ⓗ}}(j|k)$$

$$\begin{bmatrix}\hat{\underline{x}}^o(j|k)\\ \cdots \\ \underline{\sigma}^o(j|k)\end{bmatrix} - \tilde{h}^{-1}(j|k)\underline{d}'(j+1)\begin{bmatrix}\hat{\underline{x}}^o(j|k)\\ \cdots \\ \underline{\sigma}^o(j|k)\end{bmatrix} \qquad (6.4.35)$$

We have thus shown that if the solution of (6.4.31) exists and unique, the open-loop optimal control must be given by (6.4.35); and the O.L.F.O. control is given by (6.3.41). We shall consider the question of existence and uniqueness of O.L.F.O. control in the next section.

6.5 Existence and Uniqueness of O.L.F.O. Control

From equation (6.3.41), we see that if the optimal open loop control $\{u^o(j|k)\}_{j=k}^{N-1}$ exists and is unique for all $k = 0, 1, \ldots, N - 1$, we can conclude that the O.L.F.O. control $\{u^*(k)\}_{k=o}^{N-1}$ exists and is unique. If the solution of (6.3.33), i.e., the matrix $\tilde{\underline{K}}(j|k)$, exists and is unique, then the control law given by (6.3.32) and (6.3.33) is the unique globally optimal

open loop control. Since $\tilde{\underline{D}}(j|k)$ is an indefinite matrix, the solution of (6.3.33), $\tilde{\underline{K}}(j|k)$ (if it exists), is not necessary nonnegative definite; in fact it is always indefinite. Therefore, we cannot a priori conclude that $\tilde{h}(j|k) + \tilde{\underline{b}}^{o\prime}(j|k)\tilde{\underline{K}}(j+1|k)\tilde{\underline{b}}^{o}(j|k)$ will always be nonzero, and thus deduce that $\tilde{\underline{K}}(j|k)$ will remain bounded in finite time. In this section, we shall establish the existence and uniqueness of the solution of (6.3.33), $\tilde{\underline{K}}(j|k)$, for the case where the terminal time is finite ($N < \infty$); this result will then be used to prove the existence and uniqueness of the O. L. F. O. control.

Let us define

$$L^{o}(j|k) = <\underline{\hat{x}}^{o}(j|k),\underline{W}(j)\underline{\hat{x}}^{o}(j|k)>+2u^{o}(j|k)<\begin{bmatrix}\underline{\hat{x}}^{o}(j|k)\\ \cdots \\ \underline{\sigma}^{o}(j|k)\end{bmatrix},\underline{d}(j+1)>+\tilde{h}(j|k)(u^{o}(j|k))^{2} \tag{6.5.1}$$

Lemma 6.5.1: If $\tilde{h}(\ell|k) + \tilde{\underline{b}}^{o\prime}(\ell|k)\tilde{\underline{K}}(\ell+1|k)\tilde{\underline{b}}^{o}(\ell|k)$ is nonzero, $\ell = j, j+1, \ldots, N-1$, then

$$L^{o}(j|k) =<\begin{bmatrix}\underline{\hat{x}}^{o}(j|k)\\ \cdots \\ \underline{\sigma}^{o}(j|k)\end{bmatrix}, \tilde{\underline{K}}(j|k)\begin{bmatrix}\underline{\hat{x}}^{o}(j|k)\\ \cdots \\ \underline{\sigma}^{o}(j|k)\end{bmatrix}>-<\begin{bmatrix}\underline{\hat{x}}^{o}(j+1|k)\\ \cdots \\ \underline{\sigma}^{o}(j+1|k)\end{bmatrix}, \tilde{\underline{K}}(j+1|k)\begin{bmatrix}\underline{\hat{x}}^{o}(j+1|k)\\ \cdots \\ \underline{\sigma}^{o}(j+1|k)\end{bmatrix}> \tag{6.5.2}$$

Proof: Using (6.3.32), (6.3.33), (6.3.35), and (6.5.1), we have

$$L^{o}(j|k) =<\begin{bmatrix}\underline{\hat{x}}^{o}(j|k)\\ \cdots \\ \underline{\sigma}^{o}(j|k)\end{bmatrix}, \tilde{\underline{D}}(j|k)\begin{bmatrix}\underline{\hat{x}}^{o}(j|k)\\ \cdots \\ \underline{\sigma}^{o}(j|k)\end{bmatrix}>+<\begin{bmatrix}\underline{\hat{x}}^{o}(j|k)\\ \cdots \\ \underline{\sigma}^{o}(j|k)\end{bmatrix}, \underline{H}'(j|k)\tilde{\underline{K}}(j+1|k)\tilde{\underline{b}}^{o}(j|k)\cdot$$

$$(\tilde{h}(j|k)+\tilde{\underline{b}}^{o\prime}(j|k)\tilde{\underline{K}}(j+1|k)\tilde{\underline{b}}^{o}(j|k))^{-1}\tilde{h}(j|k)(\tilde{h}(j|k)+\tilde{\underline{b}}^{o\prime}(j|k)\tilde{\underline{K}}(j+1|k)\cdot$$

$$\tilde{\underline{b}}^{o\prime}(j|k)\tilde{\underline{K}}(j+1|k)\,\underline{H}(j|k)\begin{bmatrix}\underline{\hat{x}}^{o}(j|k)\\ \cdots \\ \underline{\sigma}^{o}(j|k)\end{bmatrix}> \tag{6.5.3}$$

By (6.4.34), (6.4.28), and (6.4.31), equation (6.5.3) becomes

$$L^o(j|k) = \left< \begin{bmatrix} \underline{\hat{x}}^o(j|k) \\ \cdots \\ \underline{\sigma}^o(j|k) \end{bmatrix}, \underline{\tilde{K}}(j|k) \begin{bmatrix} \underline{\hat{x}}^o(j|k) \\ \cdots \\ \underline{\sigma}^o(j|k) \end{bmatrix} \right> - \left< \begin{bmatrix} \underline{\hat{x}}^o(j+1|k) \\ \cdots \\ \underline{\sigma}^o(j+1|k) \end{bmatrix}, \underline{\tilde{K}}(j+1|k) \begin{bmatrix} \underline{\hat{x}}^o(j+1|k) \\ \cdots \\ \underline{\sigma}^o(j+1|k) \end{bmatrix} \right> \tag{6.5.4}$$

Lemma 6.5.2: For all i = k, k + 1, ..., N − 1 we have

$$\mathrm{tr}\Big\{\underline{F}\,\underline{\Sigma}_x(N|k) + \sum_{j=i}^{N-1} \underline{W}(j)\underline{\Sigma}_x(j|k)\Big\} = \mathrm{tr}\Big\{\underline{S}(i)\underline{\Sigma}_x(i|k) + \sum_{j=i}^{N-1} (\underline{S}(j+1)\underline{R}(j) +$$

$$2u(j) \left< \begin{bmatrix} \underline{x}(j|k) \\ \cdots \\ \underline{z}(j|k) \end{bmatrix}, \underline{d}(j+1) \right> + u^2(j)\underline{S}(j+1)\underline{\Sigma}_b(j)\Big\} \tag{6.5.5}$$

where u(j), j = i, i + 1, ..., N − 1, is an arbitrary control sequence and

$$\begin{bmatrix} \underline{x}(j|k) \\ \cdots \\ \underline{z}(j|k) \end{bmatrix}$$

is the resulting trajectory.

Proof: Using (6.3.28), we have ($\underline{W}(N) \triangleq \underline{F}$)

$$\mathrm{tr}(\underline{W}(j+1)\underline{\Sigma}_x(j+1|k)) = \mathrm{tr}\Big\{\underline{A}'(j)\underline{W}(j+1)\underline{A}(j)\underline{\Sigma}_x(j|k) + u(j)\underline{W}(j+1)\underline{\Sigma}_b(j|k) +$$

$$\underline{W}(j+1)\underline{R}(j) + 2u(j)\underline{W}(j+1)\underline{A}(j)\underline{\Sigma}_{xb}(j|k)\Big\} \quad ; \quad j = i, i+1, \ldots, N-1 \tag{6.5.6}$$

By applying (6.3.28) repeatedly, (6.5.6) yields

$$\mathrm{tr}(\underline{W}(j+1)\underline{\Sigma}_x(j+1|k)) = \mathrm{tr}\Big\{\underline{\phi}_A'(j,i)\underline{W}(j+1)\underline{\phi}_A(j,i)\underline{\Sigma}_x(i|k) + \sum_{\ell=i}^{j} (\underline{\phi}_A'(j,\ell+1)\underline{W}(j+1)$$

$$\underline{\phi}_A(j,\ell+1)\underline{R}(\ell) + 2u(\ell)\underline{\phi}_A'(j,\ell+1)\underline{W}(j+1)\underline{\phi}_A(j,\ell+1)\underline{A}(\ell)\underline{\Sigma}_{xb}(\ell|k) + u^2(\ell)\underline{\phi}_A'(j,\ell+1)\underline{W}(j+1)$$

$$\underline{\phi}_A(j,\ell+1)\underline{\Sigma}_b(\ell|k))\Big\} \tag{6.5.7}$$

From (6.3.19), we have for all $j \leq N-1$ that

$$\underline{S}(j) = \underline{\phi}_A'(N-1,j)\underline{F}\,\underline{\phi}_A(N-1,j) + \sum_{\ell=j}^{N-1} \underline{\phi}_A'(\ell-1,j)\underline{W}(\ell)\underline{\phi}_A(\ell-1,j) \quad . \quad (6.5.8)$$

Summing (6.5.7) over $j = i, i+1, \ldots, N-1$, and using (6.5.8) and (6.3.36), we obtain (6.5.5) after a fair amount of straightforward manipulation.

To describe the performance of the optimal open-loop control sequence, we shall introduce the notion of "conditional open loop optimal cost to go."

Definition 6.5.3: The conditional open-loop optimal cost to go for the deterministic control problem, (6.3.26)-(6.3.31);

$$J^o_{i|k}(\underline{\Sigma}(i|k),\underline{\hat{x}}(i|k)) \triangleq \min_{u(j):j=i,\ldots,N-1} \frac{1}{2}\ \underline{\hat{x}}'(N|k)\underline{F}\,\underline{\hat{x}}(N|k)+\mathrm{tr}\tilde{\underline{F}}\,\underline{\Sigma}(N|k)$$

$$+\sum_{j=i}^{N-1}[\underline{\hat{x}}'(j|k)\underline{W}(j)\underline{\hat{x}}(j|k)+\mathrm{tr}\,\tilde{\underline{W}}(j)\underline{\Sigma}(j|k)+h(j)u^2(j)] \qquad (6.5.9)$$

where $\underline{\hat{x}}(i|k)$, $\underline{\Sigma}(i|k)$ satisfy the set of equations (6.3.26)-(6.3.28).

Note that $J^o_{i|k}(\cdot,\cdot)$ is defined as a function on $M_{(2n)\times(2n)} \times R^n$. From (6.3.28), we see that $\underline{\Sigma}(j|k) \geq \underline{0}$, $j = i, i+1, \ldots, N-1$, if and only if $\underline{\Sigma}(i|k) \geq \underline{0}$. Thus, from (6.5.9) we have

$$J^o_{i|k}(\underline{\Sigma},\underline{\hat{x}}) \geq 0 \quad \text{if } \underline{\Sigma} \geq \underline{0} \quad . \qquad (6.5.10)$$

By lemma 6.5.1 and lemma 6.5.2 we immediately deduce:

Theorem 6.5.4: If $\tilde{h}(\ell|k) + \underline{\tilde{b}}^{o\prime}(\ell|k)\underline{K}(\ell+1|k)\underline{\tilde{b}}^o(\ell|k)$ is nonzero, $\ell = i, i+1, \ldots, N-1$, then the conditional open-loop optimal cost to go has the closed form

$$J^o_{i|k}(\underline{\Sigma},\underline{\hat{x}}) = \frac{1}{2}\,\mathrm{tr}\Big\{\underline{\Sigma}_x\underline{S}(i) + \sum_{j=i}^{N-1}\underline{S}(j+1)\underline{R}(j)\Big\} + \frac{1}{2}<\begin{bmatrix}\underline{\hat{x}}\\ \cdots\\ \underline{0}\end{bmatrix}, \tilde{\underline{K}}(i|k)\begin{bmatrix}\underline{\hat{x}}\\ \cdots\\ \underline{0}\end{bmatrix}>$$

(6.5.11)

where

$$\underline{\Sigma}_x = [\underline{I}_n \quad \underline{0}]\underline{\Sigma}\begin{bmatrix}\underline{I}_n \\ \cdots \\ \underline{0}\end{bmatrix} \quad ; \quad \underline{\Sigma}_{xb} = [\underline{I}_n \quad \underline{0}]\underline{\Sigma}\begin{bmatrix}\underline{0} \\ \cdots \\ \underline{I}_n\end{bmatrix} \quad ; \quad \underline{\sigma} = \begin{bmatrix}\underline{\Sigma}_{xb}\underline{e}_1 \\ \vdots \\ \underline{\Sigma}_{xb}\underline{e}_n\end{bmatrix} \quad . \quad (6.5.12)$$

We shall now make use of (6.5.11) and (6.5.10) to establish the existence, uniqueness and boundedness of $\underline{\tilde{K}}(j|k)$, $j = k, k+1, \ldots, N-1$; $k = 0, 1, \ldots, N-1$.

Lemma 6.5.5: Let $\underline{G}(j)\underline{B}\,\underline{G}'(j) \le \underline{B}$, for all $\underline{B} \ge \underline{0}$, $j = k, \ldots, N-1$. Then we have $\tilde{h}(j|k) + \underline{\tilde{b}}^{o\prime}(j|k)\underline{\tilde{K}}(j+1|k)\underline{\tilde{b}}^o(j|k) > 0$, $j = k, \ldots, N-1$.

Proof: If $\underline{F} \ge \underline{0}$, then since $h(N-1) > 0$, we have

$$\tilde{h}(N-1|k)+\underline{\tilde{b}}^{o\prime}(N-1|k)\underline{\tilde{K}}(N|k)\underline{\tilde{b}}^o(N-1|k) = h(N-1)+tr(\underline{\Sigma}_b^o(N-1|k)\underline{F})$$
$$+\underline{\hat{b}}^{o\prime}(N-1|k)\underline{F}\,\underline{\hat{b}}^o(N-1|k) > 0 \quad . \quad (6.5.13)$$

Now assume that $\tilde{h}(\ell|k) + \underline{\tilde{b}}^{\prime o}(\ell|k)\underline{\tilde{K}}(\ell+1|k)\underline{\tilde{b}}^o(\ell|k) > 0$ for $\ell = i, i+1, \ldots, N-1$ $(k < i)$. Consider the special case: $\underline{R}(j) = \underline{0}$, $j = k, \ldots, N$; then, by the induction hypothesis, theorem 6.5.4 and (6.5.10) imply that

$$J^o_{i|k}(\underline{\hat{\Sigma}}(i-1|k),\underline{\tilde{b}}^o(i-1|k)) = tr(\underline{\Sigma}_b^o(i-1|k)\underline{S}(i))+\underline{\tilde{b}}^{\prime o}(i-1|k)\underline{\tilde{K}}(i|k)\underline{\tilde{b}}^o(i-1|k) \ge 0 \quad (6.5.14)$$

where we have chosen

$$\underline{\Sigma}(i-1|k) = \begin{bmatrix}\underline{\Sigma}_b^o(i-1|k) & \vdots & \underline{\Sigma}_b^o(i-1|k)\underline{G}'(i-1) \\ \cdots & & \cdots \\ \underline{G}(i-1)\underline{\Sigma}_b(i+1|k) & \vdots & \underline{\Sigma}_b^o(i-1|k)\end{bmatrix} \ge$$
$$\begin{bmatrix}\underline{\Sigma}_b^o(i-1|k) & \vdots & \underline{\Sigma}_b^o(i-1|k)\underline{G}'(i-1) \\ \cdots & & \cdots \\ \underline{G}(i-1)\underline{\Sigma}_b^o(i-1|k) & \vdots & \underline{G}(i-1))\underline{\Sigma}_b^o(i-1|k)\underline{G}'(i-1)\end{bmatrix} \ge \underline{0} \quad (6.5.15)$$

and since $h(i-1) > 0$, we have from (6.3.37) that

$$\tilde{h}(i-1|k)+\tilde{\underline{b}}^{o\prime}(i-1|k)\tilde{\underline{K}}(i|k)\tilde{\underline{b}}^{o}(i-1|k) = h(i-1)+J^{o}_{i|k}(\hat{\underline{\Sigma}}(i-1|k),\hat{\underline{b}}^{o}(i-1|k)) > 0 \quad . \tag{6.5.16}$$

Thus the lemma is proved by induction.

We can now easily prove the existence and uniqueness of the solution of (6.3.33), $\{\tilde{\underline{K}}(j|k)\}_{j=k}^{N-1}$, $k = 0, 1, \ldots, N-1$.

Theorem 6.5.6: (Existence and Uniqueness) Let $\underline{B} \geq \underline{G}(j)\underline{B}\,\underline{G}'(j)$, $j = 0, \ldots, N-1$ for all $\underline{B} \geq \underline{0}$. The solution of (6.3.33), $\{\tilde{\underline{K}}(j|k)\}_{j=k}^{N-1}$, exists, is unique and is bounded, $(N < \infty)$, if $\underline{\Sigma}_b^{o}(k|k)$, $\hat{\underline{b}}^{o}(k|k)$, $\underline{A}(k)$, $\underline{W}(k)$, $\underline{F}\,\underline{N}(k)$ and $h(k)$ are bounded, $k = 0, 1, \ldots, N-1$.

Proof: The equation (6.3.33) can be written as a set of two equations (see chapter 2, section 2.5)

$$\tilde{\underline{K}}(j|k) = (\underline{\Theta}(j|k)-\tilde{\underline{b}}^{o}(j|k)\underline{V}_o(j+1|k))'\tilde{\underline{K}}(j+1|k)(\underline{\Theta}(j|k)-\tilde{\underline{b}}^{o}(j|k)\underline{V}_o(j+1))$$

$$+\tilde{\underline{D}}(j|k)+\underline{V}_o'(j+1)\tilde{h}(j|k)\underline{V}_o(j+1|k) \quad ; \quad \tilde{\underline{K}}(n|k) = \underline{F} \quad ; \quad j = k,k+1,\ldots,N-1$$

$$k = 0,1,\ldots,N-1 \tag{6.5.17}$$

$$\underline{V}_o(j+1|k) = (\tilde{h}(j|k)+\tilde{\underline{b}}^{o\prime}(j|k)\tilde{\underline{K}}(j+1|k)\tilde{\underline{b}}^{o}(j|k))^{-1}\tilde{\underline{b}}^{o\prime}(j|k)\tilde{\underline{K}}(j+1|k)\underline{\Theta}(j|k) \quad . \tag{6.5.18}$$

From (6.3.38), (6.3.40) and the assumption on boundedness, we have that $\tilde{\underline{b}}^{o}(j|k)$ and $\tilde{h}(j|k)$ is bounded for $j = k, k+1, \ldots, N-1$; $k = 0, 1, \ldots, N-1$. By lemma 6.5.5 $\underline{V}_o(j+1|k)$ exists, is unique and is bounded. The assertion follows from the linearity of (6.5.18) and the fact that $N < \infty$.

Corollary 6.5.7: The optimal open-loop control, $\{u^o(j|k)\}_{j=k}^{N-1}$, exists, is unique and is bounded if the assumptions in Theorem 6.5.6 are satisfied; furthermore, $\underline{\hat{x}}(k|k,U^*(0,k-1))$, $\underline{\Sigma}(k|k,U^*(0,k-1))$, $\underline{\hat{b}}(k|k,U^*(0,k-1))$ are bounded.

Theorem 6.5.8: Let $\underline{B} \geq \underline{G}(j)\underline{B}\,\underline{G}'(j)$, $j = 0, 1, \ldots, N-1$, for all $\underline{B} \geq \underline{0}$, and $\underline{A}(k)$, $\underline{B}(k)$, $\underline{X}(k)$, $\underline{G}(k)$, $\underline{W}(k)$, $h(k)$, $\underline{F}$ are bounded, $k = 0, 1, \ldots, N-1$; then the O.L.F.O. control, $u^*(k)$, $k = 0, 1, \ldots, N-1$, exists, is unique and is bounded.

Proof: We shall use induction on k. When $k = 0$, $\underline{x}(0|0)$ and $\underline{b}(0|0)$ are bounded almost surely; also $\underline{\Sigma}(0|0)$ is bounded; thus by corollary 6.5.7, $u^o(0|0)$ exists, is unique and is bounded a.s. By (6.3.41), $u^*(0)$ exists, is unique and is bounded. Assume the statement of the theorem is true for $k = 0, \ldots, i$; $i < N-1$. By the assumptions and the induction hypothesis, the identification equations, (6.3.19)-(6.3.23), imply that $\underline{\hat{x}}(i+1|i+1,U^*(0,i))$ and $\underline{b}(i+1|i+1,U^*(0,i))$ are bounded a.s., and that $\underline{\Sigma}(i+1|i+1,U^*(0,i))$ is bounded, thus corollary 6.5.7 implies that $u^o(i+1|i+1)$ exists, is unique and is bounded a.s.; by (6.3.41) the assertion of the theorem holds for $u^*(k)$, $k = 0, 1, \ldots, i+1$.

One would like to extend the results to the infinite time case with $N \to \infty$. Unfortunately, this is seldomly possible. From (6.3.39), we note that if we let $N \to \infty$, $\underline{S}(j)$ will remain bounded if and only if $\underline{A}(k)$ is exponentially stable; thus, the solution of (6.3.33), $\underline{\tilde{K}}(j|k)$, with $N \to \infty$ will not be meaningful unless $\underline{A}(k)$ is asymptotically stable. In many cases of interest, the system to be controlled is unstable. Therefore, we shall not investigate the solution of (6.3.33) with $N \to \infty$.

6.6 Asymptotic Behavior of the Identifier

In this section, we shall study the asymptotic behavior of the identification equations. The results will allow us to consider the problem of controlling the system $\underset{\sim}{S}$ over an infinite time interval ($N \to \infty$).

Definition 6.6.1: $\{(\underline{A}(k), \underline{C}(k))\}_{k=0}^{\infty}$ is said to be completely observable of index ν at k if the observation matrix

$$\underline{M}_{A,C}(k,\nu) = [\underline{C}'(k) \vdots \underline{\phi}_A'(k,k)\underline{C}'(k+1) \vdots \cdots \vdots \underline{\phi}_A'(k+\nu-2,k)\underline{C}'(k+\nu-1)] \tag{6.6.1}$$

is of full rank n. $\{(\underline{A}(k),\underline{C}(k))\}_{k=0}^{\infty}$ is said to be uniformly completely observable of index ν if the pair is completely observable of index ν for all $k = 0, 1, \ldots$.

Theorem 6.6.2: Let $\{(\underline{A}(k),\underline{C}(k))\}_{k=0}^{\infty}$ be uniformly completely observable of index ν, and suppose that $\underline{A}(k)$, $\underline{G}(k)$ are nonsingular, $k = 0, 1, \ldots$. If $u(k) \neq 0$, $k = 0, 1, \ldots$, then $\{(\underline{\tilde{A}}(k,u(k)),\underline{\tilde{C}}(k)\}_{k=0}^{\infty}$ is uniformly completely observable of index ν', $\nu' \leq 2\nu$.

Proof: By (6.3.17) and (6.3.20), we have

$$\underline{M}'_{\tilde{A},\tilde{C}}(k,2\nu) = \begin{bmatrix} \underline{C}(k) & \vdots & \underline{0} \\ \underline{C}(k+1)\underline{\phi}_A(k,k) & \vdots & \underline{C}(k+1)u(k) \\ \vdots & & \vdots \\ \underline{C}(k+j)\underline{\phi}_A(k+j-1,k) & \vdots & \sum_{\ell=k}^{k+j+1} \underline{C}(k+j)\underline{\phi}_A(k+j-1,\ell+1)u(\ell)\underline{\phi}_G(\ell-1,k) \\ \vdots & & \vdots \\ \underline{C}(k+2\nu-1)\underline{\phi}_A(k+2\nu-2,k) & \vdots & \sum_{\ell=k}^{k+2\nu-2} \underline{C}(k+2\nu-1)\underline{\phi}_A(k+2\nu-2,\ell+1)u(\ell)\underline{\phi}_G(\ell-1,k) \end{bmatrix} \tag{6.6.2}$$

By assumption, the first $m\nu$ rows of vectors contains at least n independent vectors. Among the rows vectors $\underline{C}(k+\nu+j)\underline{\phi}_A(k+\nu+j-1,k)$, let

$\underline{c}'_{\rho_j(1)}(k + \nu + j)\underline{\phi}_A(k + \nu + j - 1,k)$, ... $\underline{c}'_{\rho_j(\nu_j)}(k + \nu + j)\underline{\phi}_A(k + \nu + j - 1,k)$,

be the ν_j vectors which are independent of the row vectors:

$\underline{C}(k + \nu)\underline{\phi}_A(k + \nu - 1,k)$, $\underline{C}(k + \nu - 1)\underline{\phi}_A(k + \nu,k)$, ... $\underline{C}(k + \nu + j - 1)$ $\underline{\phi}_A(k + \nu + j - 2,k)$, $j = 1, \ldots, \nu - 1$; where

$$\underline{C}(k + \nu + j) = \begin{bmatrix} \underline{c}'_1(k + \nu + j) \\ \vdots \\ \underline{c}'_m(k + \nu + j) \end{bmatrix} \qquad (6.6.3)$$

and $\rho_j(\cdot)$ is some permutation of $\{1, 2, \ldots, m\}$. Since $\{(\underline{A}(k),\underline{C}(k))\}_{k=0}^{\infty}$ is uniformly completely observable of index ν, it follows that $\nu_i \neq 0$, $i = 1, \ldots, \nu - 1$, and that

$$m + \nu_1 + \nu_2 + \ldots \nu_{\nu-1} = n \quad . \qquad (6.6.4)$$

Assume that we have the dependence

$$\underline{C}'_{\rho_j(s)}(k+\nu+j)\underline{\phi}_A(k+\nu+j-1,k) = \sum_{i=0}^{\nu+j-1} \underline{\alpha}'_i(j,s)\underline{C}(k+i)\underline{\phi}_A(k+i-1,k); \quad 1 \leq s \leq \nu_j \quad (6.6.5)$$

where the only possible nonzero entries of $\underline{\alpha}'_i(j,s)$, $i = 0, \ldots, \nu + j - 1$, are those corresponding to independent rows of $\underline{C}(k + i)\underline{\phi}_A(k + i - 1,k)$, $i = 0, \ldots, \nu + j - 1$. If there exists no $\underline{\alpha}'_i(j,s)$, $i = 0, \ldots, \nu + j - 1$, which bears the relation (6.6.5), then the $(m(\nu + j - 1) + \rho(s))$th row vector of $\underline{M}'_{A,\tilde{C}}(k,2\nu)$ is independent of the first $m(\nu + j - 1)$ row vectors. If there exists $\underline{\alpha}'_i(j,s)$, $i = 0, \ldots, \nu + j - 1$ which gives the dependence (6.6.5), then such a dependence is unique by construction. Now assume that the $(m(\nu + j - 1) + \rho(s))$th row vector of $M'_{A,\tilde{C}}(k,2\nu)$ is dependent on the first $m(\nu + j - 1)$ row vectors, then we must also have the dependence

$$\sum_{\ell=k}^{k+\nu+j-1} \underline{C}'_{\underline{\rho}_j}(s)(k+\nu+j)\underline{\phi}_A(k+\nu+j-1,\ell+1)u(\ell)\underline{\phi}_G(\ell-1,k) =$$

$$\sum_{i=1}^{\nu+j-1} \underline{\alpha}_i'(j,s) \sum_{\ell=k}^{k+s-1} \underline{C}(k+i)\underline{\phi}_A(k+i-1,\ell+1)u(\ell)\underline{\phi}_G(\ell-1,k) \quad . \tag{6.6.6}$$

Since $\underline{A}(k)$ is nonsingular, by (6.6.5) we have

$$\sum_{i=0}^{r+j-1} \sum_{\ell=k+i}^{k+\nu+j-1} \underline{\alpha}_i'(j,s)\underline{C}(k+i)\underline{\psi}_A(k+i-1,\ell+1)u(\ell)\underline{\phi}_G(\ell-1,k) = \underline{0}_n \tag{6.6.7}$$

where

$$\underline{\psi}_A(i,j) = \underline{A}^{-1}(i)\underline{A}^{-1}(i+1) \ldots \underline{A}^{-1}(j) \quad ; \quad i > j \quad . \tag{6.6.8}$$

Since $\{(\underline{A}(k),\underline{C}(k))\}_{k=0}^{\infty}$ is uniformly completely observable, the vector

$$\underline{\alpha}^{j'}(j,s) \triangleq [\underline{\alpha}_0'(j,s) \ldots \vdots\ \underline{\alpha}_j'(j,s)] \tag{6.6.9}$$

cannot be the zero row vector, $s = 1, \ldots, \nu_j$. By assumption $\underline{G}(k)$ is nonsingular, therefore (6.6.7) is true if and only if $u(k+i) = 0$, $i = 0, 1, \ldots, j$ which is a contradiction. This result applies for $s = 1, \ldots, \nu_j$; $j = 0, 1, \ldots, \nu - 1$. Together with (6.6.4) and the remark made at the beginning of the proof, we have that $\underline{M}_{\tilde{A},\tilde{C}}(k,2\nu)$ will have rank 2n if $u(k+i) \neq 0$, $i = 0, 1, \ldots, \nu - 1$. The theorem follows from the assumption that $u(k) \neq 0$, $k = 0, 1, \ldots$.

Corollary 6.6.3: Let $\underline{A}(k)$, $\underline{G}(k)$ be bounded and nonsingular. If $\{(\underline{A}(k),\underline{C}(k))\}_{k=0}^{\infty}$ is uniformly completely observable of index ν, the error covariance matrix, $\underline{\Sigma}(k|k,U(0,k-1))$ which satisfies (6.3.21) to (6.3.23), will remain bounded for all $k = 0, 1, \ldots$ where $u(k)$ is any bounded but nonzero control for all $k = 0, 1, \ldots$.

Lemma 6.6.4: Suppose that $\underline{G}(k)$ satisfies

$$\underline{G}(k)\underline{B}\,\underline{G}'(k) \le \underline{B} \quad ; \quad \underline{B} \in M_{nn} \ , \ \underline{B} \ge \underline{0} \quad . \tag{6.6.10}$$

Let $\underline{\gamma}(k) \equiv \underline{0}$, i.e., there is no driving noise in gain dynamics; then for any control sequence, we have

$$\underline{\Sigma}_b(k+1|k+1,U(0,k)) \le \underline{\Sigma}_b(k|k,U(0,k)) \quad . \tag{6.6.11}$$

Proof: From (6.3.23) and (6.3.21), since $\underline{N}(k) = \underline{0}$, we have

$$\underline{\Sigma}_b(k+1|k+1,U(0,k)) = \underline{G}(k)\underline{\Sigma}_b(k|k,U(0,k-1))\underline{G}'(k) - [\underline{0}\vdots\underline{I}_n]\underline{V}^*(k+1|k,U(0,k))$$

$$(\tilde{\underline{C}}(k+1)\tilde{\underline{\Delta}}(k|k,U(0,k))\tilde{\underline{C}}'(k+1)+\underline{Q}(k+1))\underline{V}^{*\prime}(k+1|k,U(0,k))\begin{bmatrix}\underline{0}\\ \cdots\\ \underline{I}_n\end{bmatrix} \tag{6.6.12}$$

where $\underline{V}^*(k+1|k,U(0,k))$ satisfies (6.3.21)-(6.3.23), using (6.6.10), (6.6.11) follows immediately from (6.6.12).

An immediate consequence of lemma 6.6.4 is that if (6.6.10) is true and $\underline{\gamma}(k) \equiv \underline{0}$, then there exists $\underline{\Sigma}_b$ such that

$$\lim_{k\to\infty} \underline{\Sigma}_b(k|k,U(0,k-1)) = \underline{\Sigma}_b \quad . \tag{6.6.13}$$

Note that (6.6.13) is true independent of the observability of $\{(\underline{A}(k),\underline{C}(k))\}_{k=0}^{\infty}$. In the following theorem, we shall give sufficient conditions under which $\underline{\Sigma}_b \equiv \underline{0}$.

Theorem 6.6.5: Let $\underline{\gamma}(k) = \underline{0}$, $\underline{A}(k)$, $\underline{G}(k)$ be bounded and nonsingular and $\underline{G}(k)$ satisfies (6.6.10), $k = 0, 1, \ldots$. If $\{(\underline{A}(k),\underline{C}(k))\}_{k=0}^{\infty}$ is uniformly completely observable of index ν and $u(k)$ is any bounded but nonzero control for $k = 0, 1, \ldots$, then

$$\lim_{k\to\infty} \underline{\Sigma}_b(k|k,U(0,k-1)) = \underline{0} \quad . \tag{6.6.14}$$

Proof: Let $\varepsilon > 0$ such that

$$||\underline{\Sigma}_b(k+2\nu|k+2\nu,U(0,k+2\nu-1)) - \underline{\Sigma}_b(k|k,U(0,k-1))|| \leq \varepsilon \qquad (6.6.15)$$

where $||\cdot||$ is the spectral norm. Since $\underline{\Sigma}_b(k|k,U(0,k-1)) \geq \underline{0}$, $k = 0, 1, \ldots$, (6.6.11) and (6.6.15) imply that we have the inequality

$$||\underline{\Sigma}_b(k+j|k+j,U(0,k+j-1) - \underline{\Sigma}_b(k+j-1|k+j-1,U(0,k+j-2)|| \leq \varepsilon$$

$$j = 1, 2, \ldots, 2\nu \quad . \qquad (6.6.16)$$

Using equation (6.6.12), we have

$$\varepsilon \geq ||[\underline{0}\vdots\underline{I}_n]\underline{V}^*(k+j|k+j-1,U(0,k+j-1)\cdot(\underline{\tilde{C}}(k+j)\underline{\tilde{A}}(k+j-1|k+j-1,U(0,k+j-1)\cdot$$

$$\underline{\tilde{C}}'(k+j) + \underline{Q}(k+j))\underline{V}^{*\prime}(k+j|k+j-1,U(0,k+j-1))\begin{bmatrix}\underline{0}\\ \cdots \\ \underline{I}_n\end{bmatrix}|| \quad . \qquad (6.6.17)$$

By corollary (6.6.3), $\underline{\tilde{C}}(k+j)\underline{\tilde{A}}(k+j-1|k+j-1,U(0,k+j-1))\underline{\tilde{C}}'(k+j) + \underline{Q}(k+j)$ can be uniformly bounded, so

$$||(\underline{\tilde{C}}(k+j)\underline{\tilde{A}}(k+j-1|k+j-1,U(0,k+j-1))\underline{\tilde{C}}'(k+j)+\underline{Q}(k+j))\underline{V}^{*\prime}(k+j|k+j-1,U(0,k+j-1))\begin{bmatrix}\underline{0}\\ \cdots \\ \underline{I}_n\end{bmatrix}|| \leq$$

$$||(\underline{\tilde{C}}(k+j)\underline{\tilde{A}}(k+j-1|k+j-1,U(0,k+j-1)\underline{\tilde{C}}'(k+j)+\underline{Q}(k+j))^{\frac{1}{2}}||\cdot||(\underline{\tilde{C}}(k+j)\underline{\tilde{A}}(k+j-1|k+j-1,U(0,k+j-1))$$

$$\underline{\tilde{C}}'(k+j)+\underline{Q}(k+j))^{\frac{1}{2}}\underline{V}^*(k+j|k+j-1,U(0,k+j-1))\begin{bmatrix}\underline{0}\\ \cdots \\ \underline{I}_n\end{bmatrix}||$$

$$\leq C_j\sqrt{\varepsilon} \triangleq \delta_j(\varepsilon) \qquad j = 1, 2, \ldots, 2\nu \qquad (6.6.18)$$

$\delta_j(\varepsilon)$ is continuous in ε and $\delta_j(\varepsilon) \to 0$ as $\varepsilon \to 0$, $j = 1, \ldots, \nu$. Using (6.3.21), (6.6.18) can also be written as follows

$$\|\underline{C}(k+j)\underline{A}(k+j-1)\underline{\Sigma}_{xb}(k+j-1|k+j-1,U(0,k+j-2))\underline{G}'(k+j-1)+u(k+j-1)\underline{C}(k+j)\cdot$$

$$\underline{\Sigma}_b(k+j-1|k+j-1,U(0,k+j-2))\underline{G}'(k+j-1)\| \leq \hat{\delta}_j(\varepsilon)$$

$$j = 1, \ldots, 2\nu \quad . \tag{6.6.19}$$

Since $\underline{V}^*(k+j|k+j-1,U(0,k+j-1))$ is bounded for $j = 1, \ldots,$ therefore (6.6.17) and (6.3.21) imply that

$$\|[\underline{0} \vdots \underline{I}_n]\underline{V}^*(k+j|k+j-1,U(0,k+j-1))\tilde{\underline{C}}(k+j)\tilde{\underline{A}}(k+j-1|k+j-1,U(0,k+j-1))\begin{bmatrix}\underline{0}\\ \cdots \\ \underline{I}_n\end{bmatrix}\| \leq \hat{\delta}_1(\varepsilon) \tag{6.6.20}$$

$$\|[\underline{I}_n \vdots \underline{0}]\underline{V}^*(k+j|k+j-1,U(0,k+j-1))\tilde{\underline{C}}(k+j)\tilde{\underline{A}}(k+j-1|k+j-1,U(0,k+j-1))\begin{bmatrix}\underline{0}\\ \cdots \\ \underline{I}_n\end{bmatrix}\|$$

$$\leq \hat{\delta}_2(\varepsilon) \tag{6.6.21}$$

where $\hat{\delta}_i(\varepsilon)$ is continuous in ε, $\hat{\delta}_i(\varepsilon) \to 0$ as $\varepsilon \to 0$, $i = 1, 2, \ldots$. By using (6.3.23), (6.6.20) and (6.6.21) and the assumption that $\underline{G}(k)$ is nonsingular, the inequality (6.6.19) implies

$$\|[\underline{C}(k+1) \vdots \underline{0}]\begin{bmatrix}\underline{\Sigma}_{xb}(k+1|k+1,U(0,k))\\ \\ \underline{\Sigma}_b(k+1|k+1,U(0,k))\end{bmatrix}\| \leq f_1(\varepsilon) \tag{6.6.22}$$

$$\|\Big[\underline{C}(k+j)\underline{\Phi}_A(k+j-1,k+1) \vdots \underline{C}(k+j)\sum_{\ell=k+1}^{k+j-1}\underline{\Phi}_A(k+j-1,\ell+1)u(\ell)\underline{\Phi}_G(\ell-1,k+1)\Big]$$

$$\begin{bmatrix}\underline{\Sigma}_{xb}(k+1|k+1,U(0,k))\\ \\ \underline{\Sigma}_b(k+1|k+1,U(0,k))\end{bmatrix}\| \leq f_j(\varepsilon)$$

$$j = 2, 3, \ldots, 2\nu \tag{6.6.23}$$

where $f_i(\varepsilon)$ is continuous in ε, $f_i(\varepsilon) \to 0$ as $\varepsilon \to 0$, $i = 1, 2, \ldots, 2\nu$. Equations (6.6.22) and (6.6.23) imply

$$\left|\left| \underline{M}_{\tilde{A},\tilde{C}}(k+1,2\nu)\begin{bmatrix} \underline{\Sigma}_{xb}(k+1|k+1,U(0,k)) \\ \\ \underline{\Sigma}_b(k+1|k+1,U(0,k)) \end{bmatrix}\right|\right| \le f(\varepsilon) \tag{6.6.24}$$

where $f(\varepsilon) \to 0$ when $\varepsilon \to 0$ and is continuous in ε. By theorem 6.6.2, $\underline{M}_{\tilde{A},\tilde{C}}(k + 1,2\nu)$ is of full rank, so we have

$$||\underline{\Sigma}_{xb}(k+1|k+1,U(0,k))|| \le \delta'(\varepsilon) \qquad \delta'(\varepsilon) \to 0 \qquad \text{as } \varepsilon \to 0 \tag{6.6.25}$$

$$||\underline{\Sigma}_b(k+1|k+1,U(0,k))|| \le \delta''(\varepsilon) \qquad \delta''(\varepsilon) \to 0 \qquad \text{as } \varepsilon \to 0 \quad . \tag{6.6.26}$$

Now the conclusion of the theorem follows from (6.6.13).

Theorem 6.6.5 can be extended to the case where u(k) is bounded but nonzero control for all but a finite number of k's. Since $\underline{\Sigma}(k|k,U(0,k-1)) \ge \underline{0}$, (6.6.14) also implies

$$\lim_{k\to\infty} \underline{\Sigma}_{xb}(k|k,U(0,k-1)) \to \underline{0} \tag{6.6.27}$$

if the conditions for theorem 6.6.5 hold.

Let us consider an observable system S, (6.2.1), the gain parameters are assumed to be unknown and satisfy

$$\underline{b}(k + 1) = \underline{G}(k)\underline{b}(k) \tag{6.6.28}$$

with $\underline{G}(k)$ satisfying (6.6.10). Assume that we want to control the system S over an interval $N < \infty$. In the beginning, the modified weighting on the control is high, and thus in general, the control magnitude will be low at

the beginning. Thus, the trajectory of the overall control system would be pretty much the same as the input-free trajectory of the system $\mathcal{S}$. If the matrix $\underline{A}(k)$ is exponentially stable, the true state of the system will evolve toward zero by using negligibly small control magnitudes (even zero). The result is that little effort of the input, $\{u(k)\}_{k=0}^{N-1}$ is spent for control and identification purposes. We would expect that the estimated parameters will hardly converge to the true parameters, $\underline{b}(k)$. On the other hand if $\underline{A}(k)$ is not exponentially stable, then the true state of the overall system will diverge. This diverging phenomenon will be noticed by the identifier, thus resulting in a high control magnitude because of (6.3.32). Since little is initially known about the gain parameters, the high magnitude control will be utilized mainly for identification purposes. Therefore the control will be kept bounded away from zero as long as exact identification of $\underline{b}(k)$ has not been obtained. Using theorem 6.6.5, we predict that the estimated parameters of $\underline{b}(k)$ will converge to the true gain parameters before the control magnitude goes to zero.

Analytical studies of the convergence rate of the O.L.F.O. system are not yet available. From the above discussion, we may predict roughly that the convergence-rate for unstable system will be relatively fast depending on "how stable" the system is; and the convergence-rate for stable system will be very slow.

For control over an infinite time period, see section 6.7(C) for detailed discussions.

Finally, we shall discuss some interesting implications of theorem 6.6.5. Consider an observable system $\mathcal{S}$, (6.2.1), with unknown gain

parameters satisfying (6.6.28) and with $\underline{G}(k)$ satisfies (6.6.10). Let $\underline{c}_k(\hat{\underline{x}}(k|k,\hat{\underline{b}}(k|k),\underline{\Sigma}_b(k|k,U(0,k-1))$ be any ad-hoc control law which is "put" after the identifier (see Figure 6.4) and with the following properties $(k \geq 0)$:

1) $\underline{c}_k(\cdot,\cdot,\cdot): R^n \times R^n \times M_{nn} \to R$

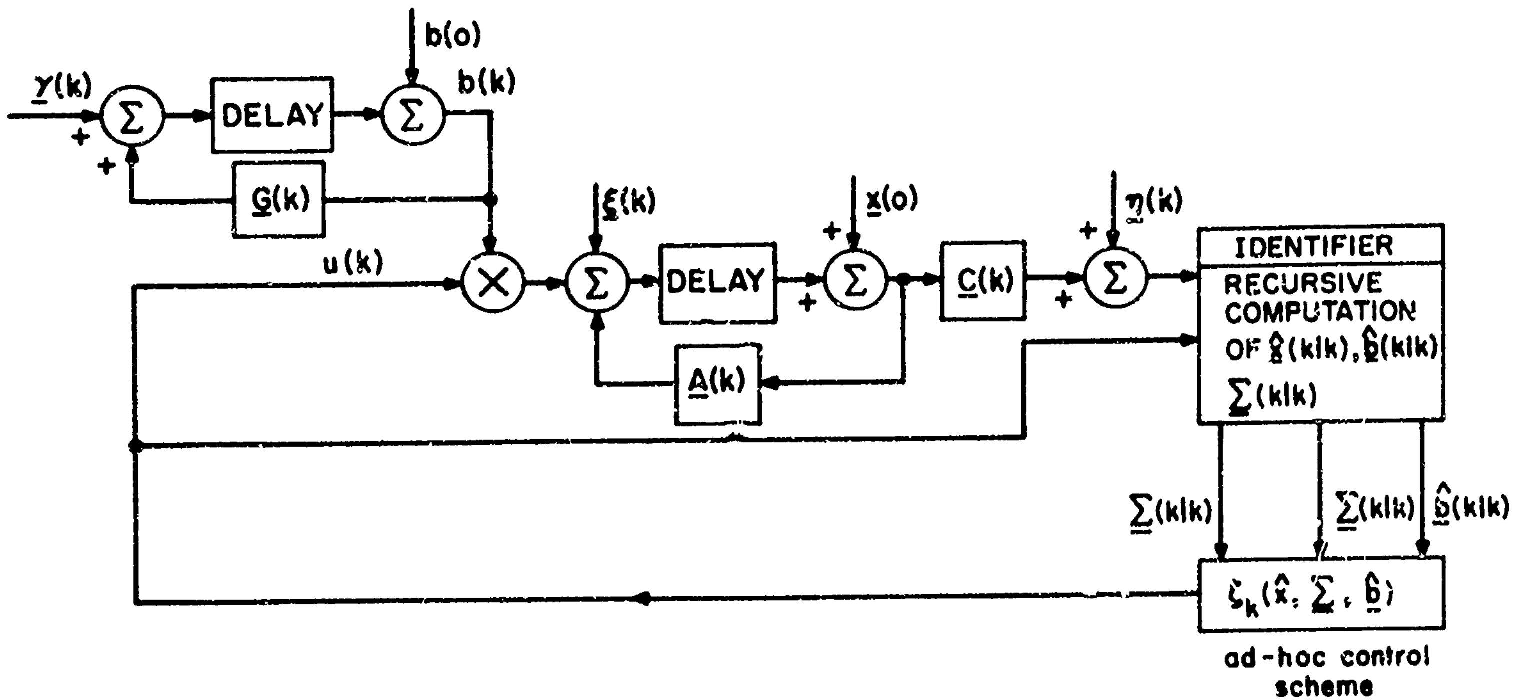
2) $\underline{c}_k(\underline{x},\underline{b},\underline{\Sigma}) \neq \underline{0}, \quad \underline{x} \in R^n,\ \underline{b} \in R^n,\ \underline{\Sigma} \in M_{nn},\ \underline{x} \neq \underline{0};\ \underline{\Sigma} \neq \underline{0}$

3) $\underline{c}_k(\underline{x},\underline{b},\underline{0}) = -(h(k)+\underline{b}'(k)\underline{K}(k+1)\underline{b}(k))^{-1}\underline{b}'(k)\underline{K}(k+1)\underline{A}(k)\underline{x}$;
$\underline{x} \in R^n,\ \underline{b} \in R^n$

From condition 2, we see that $\underline{\Sigma}_b(k|k,U(0,k-1)) \to \underline{0}$ as $k \to \infty$ (6.6.14); and so from condition 3, the ad-hoc control scheme will converge to the optimal control strategy when the full dynamics become known. This indicates that the ad-hoc scheme $\underline{c}_k(\hat{\underline{x}}(k|k),\hat{\underline{b}}(k|k),\underline{\Sigma}_b(k|k,U(0,k-1))$ can provide reasonable simulation results.

6.7 General Discussion

In this chapter, we investigated the problem of identification and control of discrete linear systems with unknown gain from the theoretical standpoint. The control is open-loop feedback optimal. The implementation of such a control (O.L.F.O. control) was described to some detail. The actual implementation for O.L.F.O. control for third-order systems will be discussed in more detail in chapter 7. As we shall see later, such a proposed scheme appears to be computationally feasible and that the results are reasonable and appealing. A deeper theoretical understanding of the derived O.L.F.O. control is possible from the results in sections 6.4 and 6.5. The questions of existence, uniqueness of O.L.F.O. control are considered in great detail. The asymptotic behavior of such control systems was treated in section 6.6; some of its extensions will be discussed later in this section.



$$\zeta_k(\hat{\underline{x}}, \underline{\Sigma}, \hat{\underline{b}}) \neq 0 \text{ if } \Sigma \neq \underline{0}$$

$$\zeta_k(\hat{\underline{x}}, \underline{0}, \hat{\underline{b}}) = -(h(k) + \underline{b}'\underline{K}(k+1)\underline{b})^{-1} \cdot \underline{b}'\underline{K}(k+1)\underline{A}(k)\hat{\underline{x}}$$

Fig. 6.4 BLOCK DIAGRAM FOR AD-HOC CONTROL SYSTEM

(A) Discussion on Approaches

The problem of combining identification and control of linear system with unknown gain have been considered by several people. Farison [60] considered an ad-hoc procedure which basically assumes the separation between identification and control. Murphy [61] considered the approximate effect of iteration between control and identification and he pre-supposed that the control was a pure feedback of the estimated states. Gorman and Zaborszky [62] used a similar approach to that of Murphy and obtained a suboptimal control which required the solution of a sequence of two point boundary value problems. Essentially, [61] and [62] are approximately Bellman's equation. The approach taken in this chapter is different from those in [60], [61], and [62].

Bar-Shalom and Sivan† [63] also used the O.L.F.O. control approach to consider control problems with random parameters. They derived a general solution but made no attempt to study analytically the derived results. The approach taken in this chapter is primarily motivated by computational feasibility

From the discussion made at the end of section 6.6, we can see why different computation schemes suggested by Farison, Murphy, Gorman and Zaborszky will all be expected to give reasonable simulation results. It is hard to quantitatively compare our approach with theirs without extensive simulation experiments. One computation advantage of our results over those

---

†This reference was brought to the author's attention when most of the theoretical work of this chapter (sections 6.2 to 6.5) had been completed.

of Murphy, Gorman and Zaborszky is that we replace a sequence of two point boundary problems by solving a sequence of matrix Riccati type difference equation of dimension $(n+1)n \times (n+1)n$, (6.3.33). This matrix Riccati difference equation is solved backward in time starting from the terminal time to the present time k; i.e., an $N - k$ steps computation, where $k = 0, 1, \ldots, N - 1$. Computational wise, this is easier than solving a two point boundary problem. In our approach, the theoretical proof on the existence and uniqueness of O.L.F.O. control sequence is available; this gives us confidence in trying out the suboptimal control scheme using a digital computer. Also, we can deduce and predict roughly the behavior of the overall O.L.F.O. control system (section 6.6) from the derived equations (section 6.3).

(B) Vector Control

In our investigation, we assumed that the control is scalar. However, the approach can be extended in a straightforward conceptual manner to the vector control case. First, a set of identification equation is derived which will generate the estimate of the current state, the current estimate of the unknown gain matrix and the different cross-error-covariance matrices. An open-loop control problem is formulated as in section 6.3, equations (6.3.20) to (6.3.31) and discrete matrix minimum principle is used to obtain the extremal solution. The results will be similar to those of scalar control case. However, the equations in the vector control case will be more complicated.

(C) Control Over Infinite Interval

Let us consider the problem of controlling the system S, which is time invariant and unknown constant gain $\underline{b}$, over an infinite interval, i.e., $N \to \infty$.

It was pointed out (in section 6.5) that the problem will not be very meaningful in many cases if we just consider the obtained results (section 6.3) and let $N \to \infty$. We suggest the <u>window-shifting</u> approach. Assume that at all times, we have N more steps to control (see Figure 6.5), thus at all times we solve an open-loop control problem over an interval of N steps. This approach is motivated by computational consideration and the theoretical results derived in section 6.6.

We note that in the O.L.F.O. approach, we have to resolve the open-loop control problem at all time k so as to adjust the control scheme accordingly. In our case, we have to compute $\tilde{\underline{K}}(k|k)$ in a backward direction starting from the terminal time N to k for each k. If N is very large, this computation will require a very long time to accomplish. From a computational standpoint, we would like to "cut back" the terminal time. Conceptually, in trying to control over an infinite time period, the controller looks into all future effects caused by present action, and decides on the optimum move. The window-shifting approach suggests that instead of looking at <u>all</u> future effects, the controller looks at only <u>near</u> future effects caused by present actions and decides on suboptimal moves. One may view such an approach as a "short term adaptive scheme." Note also that we can adjust the "window width" according to computational capability. At all times, we need only to solve for $\tilde{\underline{K}}(k|k)$ in a backward direction starting from $N + k$ to k. Thus from a conceptual and a computational point of view, such an approach may be desirable.

Assume that the time invariant system S being controlled is observable and controllable. If $\underline{b}$ is known exactly, then if we consider control over infinite time period, the optimal feedback gain is constant and is given by

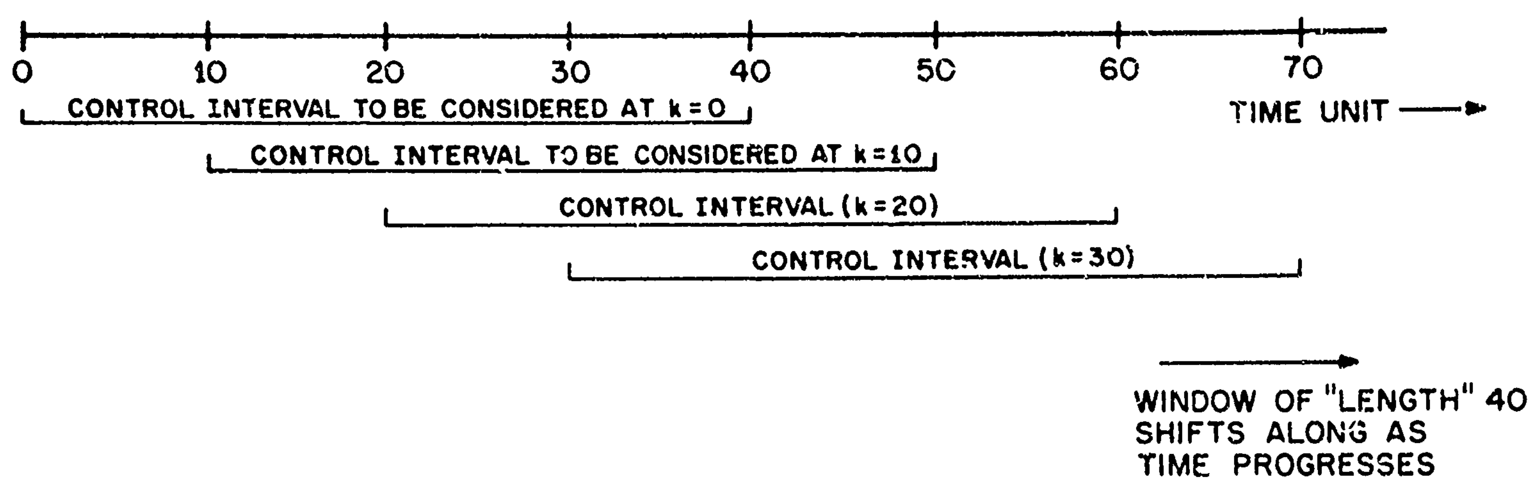


Fig. 6.5 DESCRIPTION OF WINDOW SHIFTING APPROACH.
THE "WIDTH" OF THE WINDOW IS N = 40 STEPS.

$$\underline{c} = -(h + \underline{b}'\underline{K}\,\underline{b})^{-1}\underline{b}'\underline{K}\,\underline{A} \tag{6.7.1}$$

where $\underline{K}$ is given by the steady state solution of

$$\underline{K}_{i+1} = \underline{A}'(\underline{K}_i - \underline{K}_i\underline{b}(h + \underline{b}'\underline{K}_i\underline{b})^{-1}\underline{b}'\underline{K}_i)\underline{A} + \underline{W} \quad ; \quad \underline{K}_o = \underline{F} \tag{6.7.2}$$

(see chapters 3 and 4). Let N be the integer such that for $n \geq N$,

$$||\underline{K}_n - \underline{K}_{n-1}|| \leq \varepsilon \quad ; \quad \varepsilon > 0 \quad . \tag{6.7.3}$$

Such an integer N can be found experimentally off-line. Adjust the window width to equal to N, and apply the window-shifting approach. Add some nonzero control for identification purpose if it is necessary (see also chapter 7). Using the results in section 6.5, the existence and uniqueness of such control sequence is guaranteed. By theorem 6.6.5, the estimate in $\underline{b}$ will converge asymptotically, and so when $\hat{\underline{b}}(k|k,U^*(0,k-1) \to \underline{b}$, we have

$$\tilde{\underline{K}}(k|k) \to \begin{bmatrix} \underline{K}(k,N+k;\underline{F}) & \vdots & \underline{0} \\ \cdots & & \cdots \\ \underline{0} & \vdots & \underline{0} \end{bmatrix}$$

where $\underline{K}(k,N+k;\underline{F})$ satisfies

$$\underline{K}(k,N+k;\underline{F}) = \underline{A}'(\underline{K}(k+1,N+k;\underline{F}) - \underline{K}(k+1,N+k;\underline{F})\underline{b}(h + \underline{b}'\underline{K}(k+1,N+k;\underline{F})\underline{b})^{-1}\cdot$$

$$\underline{b}'\underline{K}(k+1,N+k;\underline{F}))\underline{A} + \underline{W} \quad ; \quad \underline{K}(N+k,N+k;\underline{F}) = \underline{F} \tag{6.7.4}$$

and

$$u^*(k|k) \to \underline{\phi}(k)\hat{\underline{x}}^o(k|k) = -(h + \underline{b}'\underline{K}(k,N+k;\underline{F})\underline{b})^{-1}\underline{b}'\underline{K}(k,N+k;\underline{F})\underline{A}\,\hat{\underline{x}}^o(k|k) \tag{6.7.5}$$

(See discussion at the end of section 6.3.) Comparing (6.7.2) and (6.7.4), we note that

$$\underline{K}(k,N+k;\underline{F}) = \underline{K}_N = \underline{K} \quad . \tag{6.7.6}$$

Thus asymptotically, we have a time invariant overall control system.

(D) Convergence-Rate

We have not studied in detail (analytically) the convergence-rate of the suboptimal O.L.F.O. control system. We can only deduce and predict some rough qualitative estimates about convergence-rate for stable and unstable systems. We shall study the question of convergence-rate via simulations; some conclusions and discussion will be included in the next chapter.

(E) Conditions for Convergence[+]

From theorem 6.6.5, we note that if $\underline{\gamma}(k) = \underline{0}$, the sufficient conditions for convergence are observability, nonzero control and (6.6.10). The first two conditions are relatively easy to understand and intuitively appealing. The third condition needs some explanation.

Suppose that $\underline{G}(k)$ satisfies (6.6.10); then by taking $\underline{B} = \underline{I}_n$, we have

$$\underline{G}(k)\underline{G}'(k) \leq \underline{I}_n \quad . \tag{6.7.7}$$

Thus, we have

$$||\underline{G}(k)|| \leq 1 \tag{6.7.8}$$

where $||\cdot||$ is the spectral norm. Equation (6.7.8) provides us with the necessary condition for (6.6.10) to hold. Intuitively, (6.7.8) means that the uncertainty of $\underline{b}(k)$ cannot grow.

Let $\underline{G}(k)$ be an nxn matrix such that

$$\underline{G}(k)\underline{x} \leq \underline{x} \quad ; \quad \underline{x} \in R^n \quad . \tag{6.7.9}$$

[+]This discussion was motivated from a suggestion made by Prof. J. C. Willems.

If $\underline{B} \geq \underline{0}$, then $\underline{B} = \underline{DD}'$ for some $\underline{D} \in M_{nm}$. Let $\underline{D} = [\underline{d}_1 \ldots \underline{d}_m]$, equation (6.7.9) implies

$$\underline{G}(k)\underline{DD}'\underline{G}'(k) \leq \underline{DD}' \qquad (6.7.10)$$

and so $\underline{G}(k)$ satisfies (6.6.10). Thus (6.7.9) provides a sufficiency test for (6.6.10). Geometrically, (6.7.9) implies that $\underline{G}(k)$ is a linear transformation which is directionally invariant but shrinking or retaining the length of each vector. Some weaker sufficiency tests which have some physical interpretations will be explored in future research efforts.

(F) Different Cost Criterion

The approach can be applied to the more general case where a cost criterion other than quadratic is being considered. The identification equations remain unchanged but the open-loop control problem thus formulated will be different from (6.3.26)-(6.3.31). By using the discrete matrix minimum principle, we shall obtain a set of equations which define a two point boundary values problem.

6.8 Perspective

The problem of stochastic control of linear systems with unknown gain was also treated by Florentin [64], Farison [60], Murphy [61], Gorman and Zaborszky [62]. The approach in [61]-[64] is that of approximating the solution of Bellman's equation. [60] presuppose separation.

Open-loop feedback controller was described by Dreyfus [26]. Open-loop feedback optimal control approach was also used by Bar-Shalom and Sivan [63] in considering control of discrete-time linear systems with random parameters.

To the author's knowledge, for this particular problem of controlling linear system with unknown gain, the investigations in sections 6.5 and 6.6 represent the first extensive analytical studies on the derived suboptimal solution. The contributions being that a plausible computationally feasible suboptimal solution is derived using the O.L.F.O. approach, extensive analytical studies on the derived solution are carried out, and from the derived results some rough behavior of the overall suboptimal control system can be deduced; also we have a deeper understanding on the effects (qualitatively) of uncertainties on the control action.

CHAPTER VII

CONTROL OF THIRD ORDER SYSTEMS WITH UNKNOWN ZEROES: NUMERICAL EXAMPLES

In the last chapter, we have studied theoretically the problem of control of a discrete time linear system with unknown gain under the quadratic criterion. A suboptimal adaptive control system was derived using the O.L.F.O. approach, and the asymptotic behavior of the control system was discussed. There are still some important questions which have not been treated theoretically. For example, the rate of convergence of the suboptimal control system is in general of great interest, but was not treated in detail. Computer studies were carried out on some specific examples of third order systems. The main purpose for these studies is to provide us with some qualitative ideas about the rate of convergence of the suboptimal control system for different types of third order plants.

Let us consider a stochastic continuous time-invariant linear system described by:

$$\begin{aligned} \underline{\dot{x}}_f(t) &= \underline{A}_f\, \underline{x}_f(t) + \underline{b}_f\, u_f(t) + \underline{d}_f\, \xi_f(t) \quad ; \quad \underline{x}(0) \sim G(\underline{0}, \underline{\Sigma}_{xo}) \\ \underline{y}_f(t) &= \underline{c}'\, \underline{x}_f(t) + \eta_f(t) \qquad\qquad\qquad\quad \underline{b} \sim G(\underline{0}, \underline{\Sigma}_{bo}) \end{aligned} \tag{7.1}$$

where $\xi_f(t)$ is a scalar driving white Gaussian noise, $\eta_f(t)$ is the scalar observation white Gaussian noise. The statistical laws of $\xi_f(t)$ and $\eta_f(t)$ are assumed to be known:

$$\int_{t_1}^{t_2} \xi_f(t)dt \sim G\left(0, \int_{t_1}^{t_2} r\, dt\right) \tag{7.2}$$

$$\int_{t_1}^{t_2} \eta_f(t)dt \sim G\left(0, \int_{t_1}^{t_2} q\, dt\right) \tag{7.3}$$

From (7.1), we have

$$\underline{x}_f(t) = e^{\underline{A}_f t}\,\underline{x}_f(0) + \int_0^t e^{\underline{A}_f(t-\tau)}\,\underline{b}_f u_f(\tau)\,d\tau + \int_0^t r^{\underline{A}_f(t-\tau)}\,\underline{d}_f \xi_f(\tau)d\tau \quad (7.4)$$

Assume that we take observations only at discrete instants of time $t = \Delta$, 2Δ, 3Δ ...; Δ is assumed to be small such that $u(t) = u(k\Delta)$, $\xi(t) = \xi(k\Delta)$, $t \in [k\Delta, (k+1)\Delta]$:

$$\begin{aligned}\underline{x}_f\ (k+1)\Delta &= e^{\underline{A}_f\Delta}\Big[e^{\underline{A}_f(k\Delta)}\,\underline{x}_f(0) + \int_0^{k\Delta} e^{\underline{A}_f(k\Delta-\tau)}\,\underline{b}_f\ u_f(\tau)d\tau \\ &+ \int_0^{k\Delta} e^{\underline{A}_f(k\Delta-\tau)}\,\underline{d}_f\ \xi_f(\tau)d\tau\Big] \\ &+ \int_0^{\Delta} e^{\underline{A}_f\sigma}d\sigma\ \cdot \underline{b}_f\ u_f(k\Delta) + \int_0^{\Delta} e^{\underline{A}_f\sigma}d\sigma\ \cdot \underline{d}_f\ \xi_f(k\Delta) \end{aligned} \quad (7.5)$$

Defining

$$\begin{aligned}&\underline{x}(k) = \underline{x}_f(k\Delta)\ ;\ \underline{A} = e^{\underline{A}_f\Delta}\ ;\ \underline{b} = \int_0^{\Delta} e^{\underline{A}_f\sigma}\,d\sigma\cdot\underline{b}_f \\ &\underline{d} = \int_0^{\Delta} e^{\underline{A}_f\sigma}\,d\sigma\ \cdot\underline{d}_f\ ;\ \xi(k) = \xi_f(k\Delta)\ ;\ u(k) = u_f(k\Delta)\end{aligned} \quad (7.6)$$

(7.5) becomes

$$\underline{x}(k+1) = \underline{A}\ \underline{x}(k) + \underline{b}\ u(k) + \underline{d}\xi\ (k)\ ;\ \underline{x}(0) \sim G(\underline{0}, \underline{\Sigma}_{xo}) \quad (7.7)$$

Defining

$$y(k) = y_f(k\Delta)\ ;\ \eta(k) = \eta_f(k\Delta) \quad (7.8)$$

The observation sequence is

$$y(k) = \underline{c}'\underline{x}(k) + \eta(k) \quad (7.9)$$

The statistical laws of $\xi(k)$, $\eta(k)$ are

$$\xi(k) \sim G(0, r\Delta) \quad (7.10)$$

$$\eta(k) \sim \mathcal{G}(0,\ q\Delta) \tag{7.11}$$

The gain vector is assumed to be unknown but constant, therefore the equation for the unknown gain is

$$\underline{b}(k+1) = \underline{b}(k) \quad ; \quad \underline{b}(0) \sim \mathcal{G}(\underline{b}_o,\ \underline{\Sigma}_{bo}) \tag{7.12}$$

We can now apply the results in chapter 6 to equations (7.6), (7.9), (7.10)-(7.12).

A computer program was designed which operates as follows:

(1) Read in $\underline{A}_f$, $\underline{b}_f$, $\underline{c}$, $\underline{d}_f$, r, q, , $\underline{x}_o$, $\underline{b}_c$, the final time N and the different weightings $\underline{W}$, h, $\underline{F}$, and covariances $\underline{\Sigma}_{xo}$, $\underline{\Sigma}_{bo}$.

(2) A subroutine, which was developed by Levis [75], was used to convert the continuous version, (7.1), to the discrete time sample data version (7.6). The covariances of $\xi(k)$, $\eta(k)$ are computed using (7.10), (7.11).

(3) The true value of $\underline{x}(k)$ was recorded. Using a noise generating subroutine, a sample value of $\underline{y}(k)$ was obtained. Assume that $\hat{\underline{x}}(k-1/k-1)$, $\hat{\underline{b}}(k-1/k-1)$ are recorded. A subroutine for the identification equations (6.3.19)-(6.3.23) was used to obtain the current estimates $\hat{\underline{x}}(k/k)$, $\hat{\underline{b}}(k/k)$, and the error covariance matrix $\underline{\Sigma}(k/k)$ recursively. These values were also recorded.

(4) A subroutine based on (6.3.32)-(6.3.41) was used to obtain the adaptive control $u^*(k)$.

(5) The control $u^*(k)$ was applied to the system (7.6), using a noise generating device to obtain a sample value of $\xi(k)$; then by (7.6), we obtained the value $\underline{x}(k+1)$.

(6) We advance $k \to k+1$ and repeat (3) through (5) until we get to the final time $k = N-1$.

The program was written in such a way that if we set $\hat{\underline{b}}(k/k) = \underline{b}$, and $\underline{\Sigma}_{bo} = 0$, then the procedures (3) through (6) will give us the truly optimal stochastic control when $\underline{b}$ is known. Using a plotting subroutine we can plot out the truly optimal trajectories vs. the O.L.F.O. trajectories; the true $\underline{b}$ vs. the estimated $\underline{b}$, and optimal feedback gain vs. adaptive gain (it was noted that the adaptive correction term will converge to zero quite fast), under the requirement that the same noise samples $(\xi(k), \eta(k))$ were used for both the known $\underline{b}$ and unknown $\underline{b}$ cases. These plots provide us with qualitative understanding on the rate of convergence of the overall suboptimal O.L.F.O. control system.

In all the computer simulations, unless otherwise mentioned, we set the values:

$$\Delta = 0.2 \text{ sec}, \quad r = 0.05, \quad q = 0.45, \quad \underline{d}_f = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \underline{x}_o = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$
$$\underline{F} = \underline{I}_3, \quad \underline{W} = \underline{I}_3, \quad \underline{\Sigma}_{bo} = \underline{\Sigma}_{xo} = 4\,\underline{I}_3, \quad \underline{c}' = [1\ 0\ 0] \tag{7.13}$$

Example 1: Unstable System

It is assumed that

$$\underline{A}_f = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 5 & -3 & -1 \end{bmatrix} \quad ; \quad \underline{b}_f = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix} \quad ; \quad h = 1 \quad ; \quad \underline{x}_f(0) = \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix} \tag{7.14}$$

such a system has a transfer function (see Fig. 7.1)

$$H_1(s) = \frac{(s+3)(s+2)}{(s-1)(s^2+2s+s)} \tag{7.15}$$

so that it has an unstable pole at $s = 1$. Initially, we set

$$\hat{\underline{b}}_f(0/0) = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \tag{7.16}$$

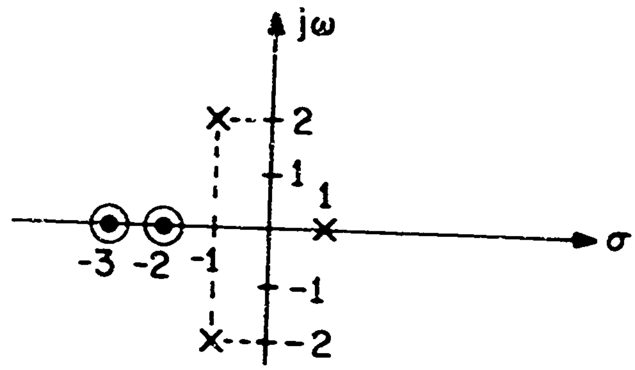


Fig. 7.1a POLE ZERO PATTERN FOR EXAMPLE 1: UNSTABLE SYSTEM

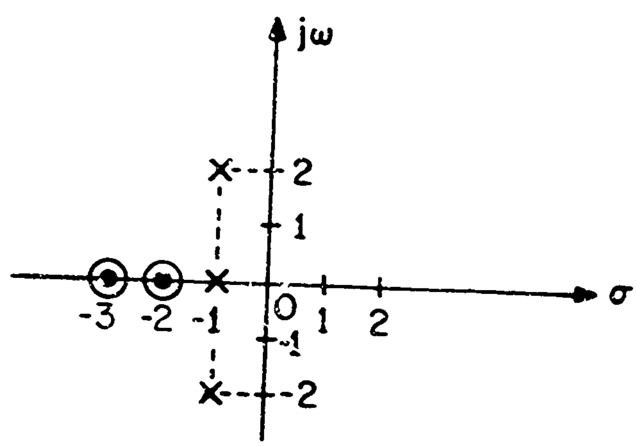


Fig. 7.1b POLE ZERO PATTERN FOR EXAMPLE 2: STABLE SYSTEM

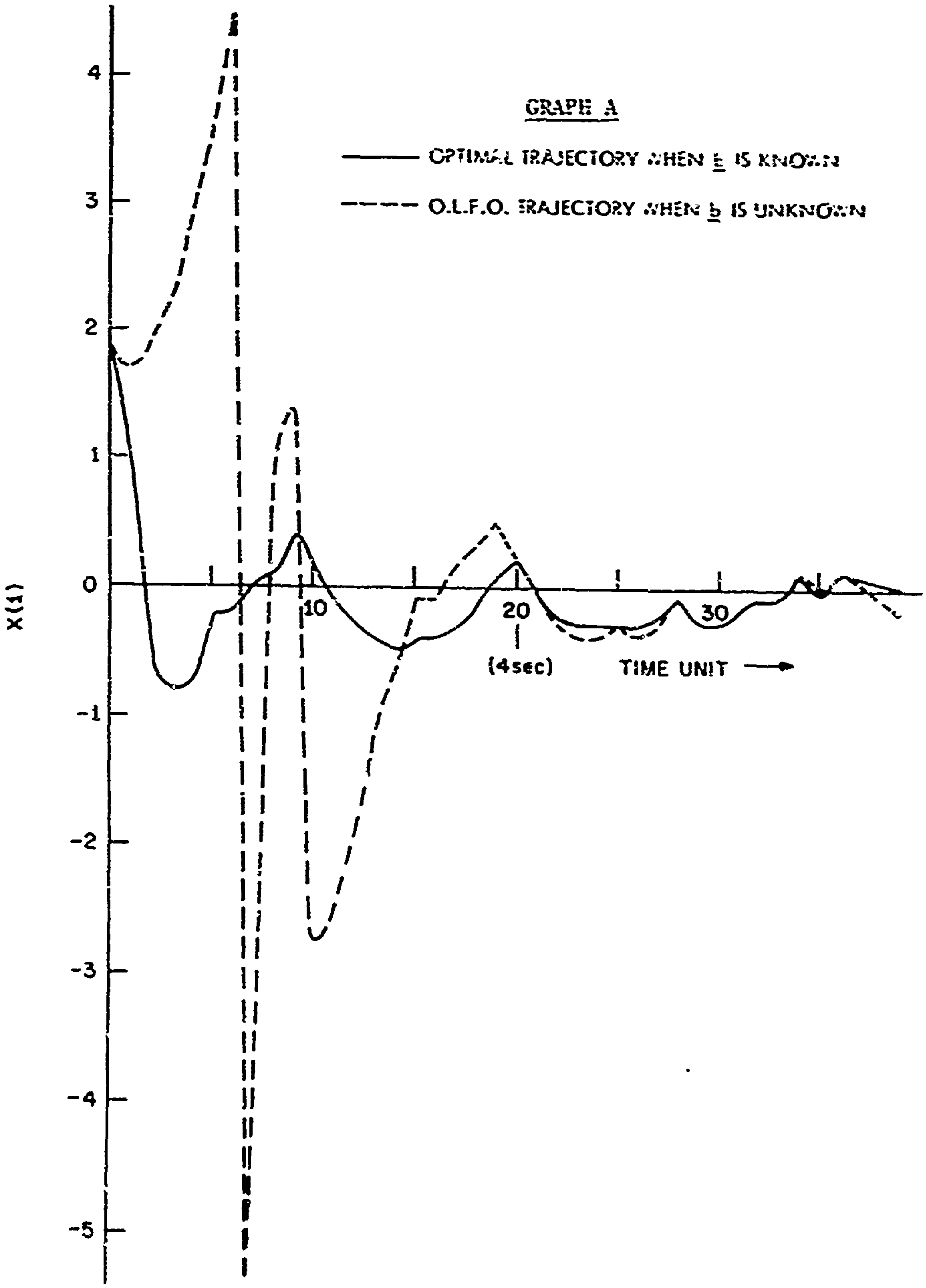


Fig. 7.2 COMPARISON BETWEEN THE OPTIMAL TRAJECTORY WHEN THE GAIN IS KNOWN AND THE O.L.F.O. TRAJECTORY ASSUMING THE GAIN IS UNKNOWN. THE SYSTEM BEING CONTROLLED IS UNSTABLE WITH SYSTEM FUNCTION $\frac{(S+3)\ (S+2)}{(S-1)\ (S^2+2S+5)}$.

THE SAMPLE NOISE IS THE SAME FOR BOTH CASES.

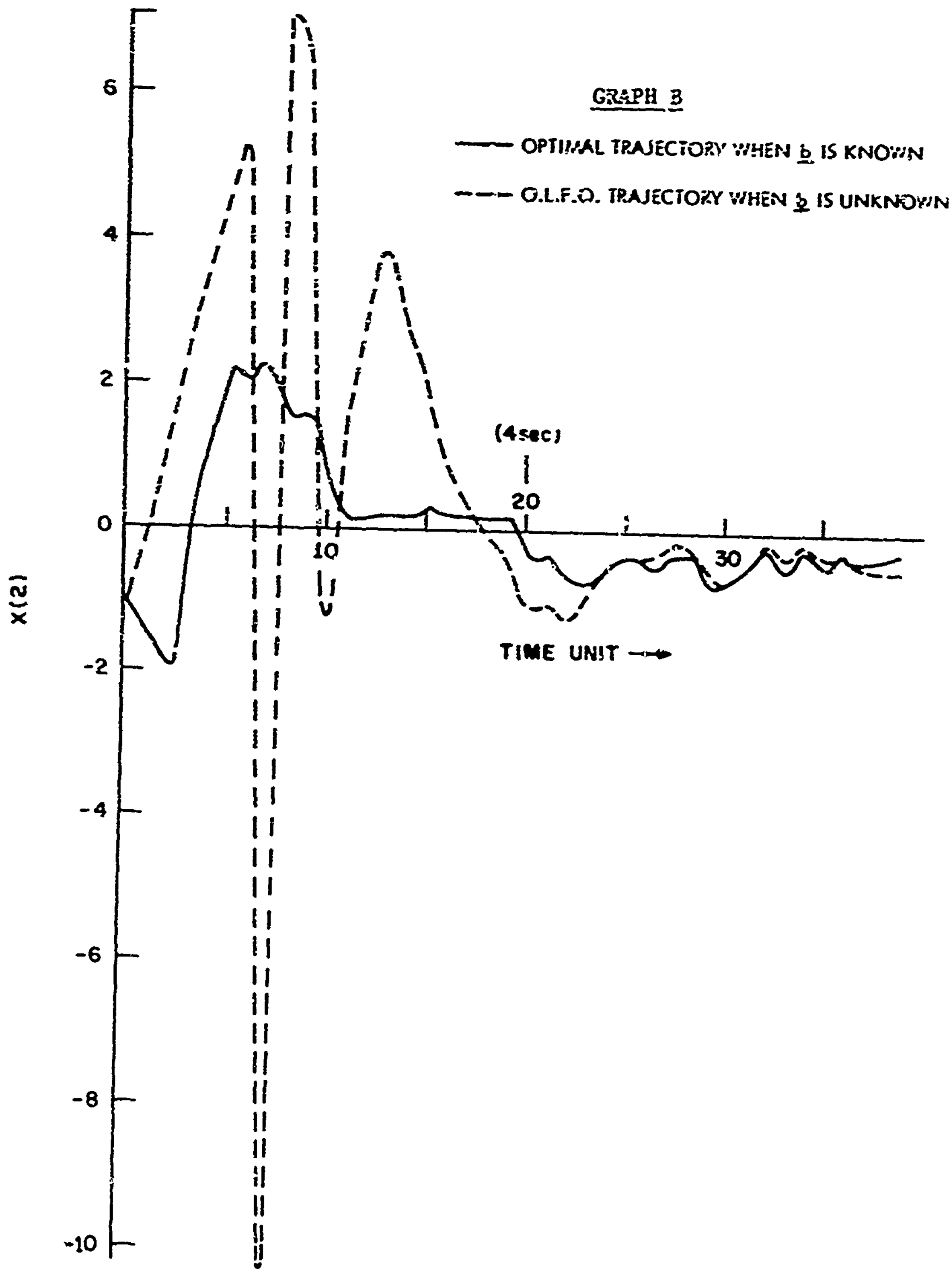


Fig. 7.2 (Continued)

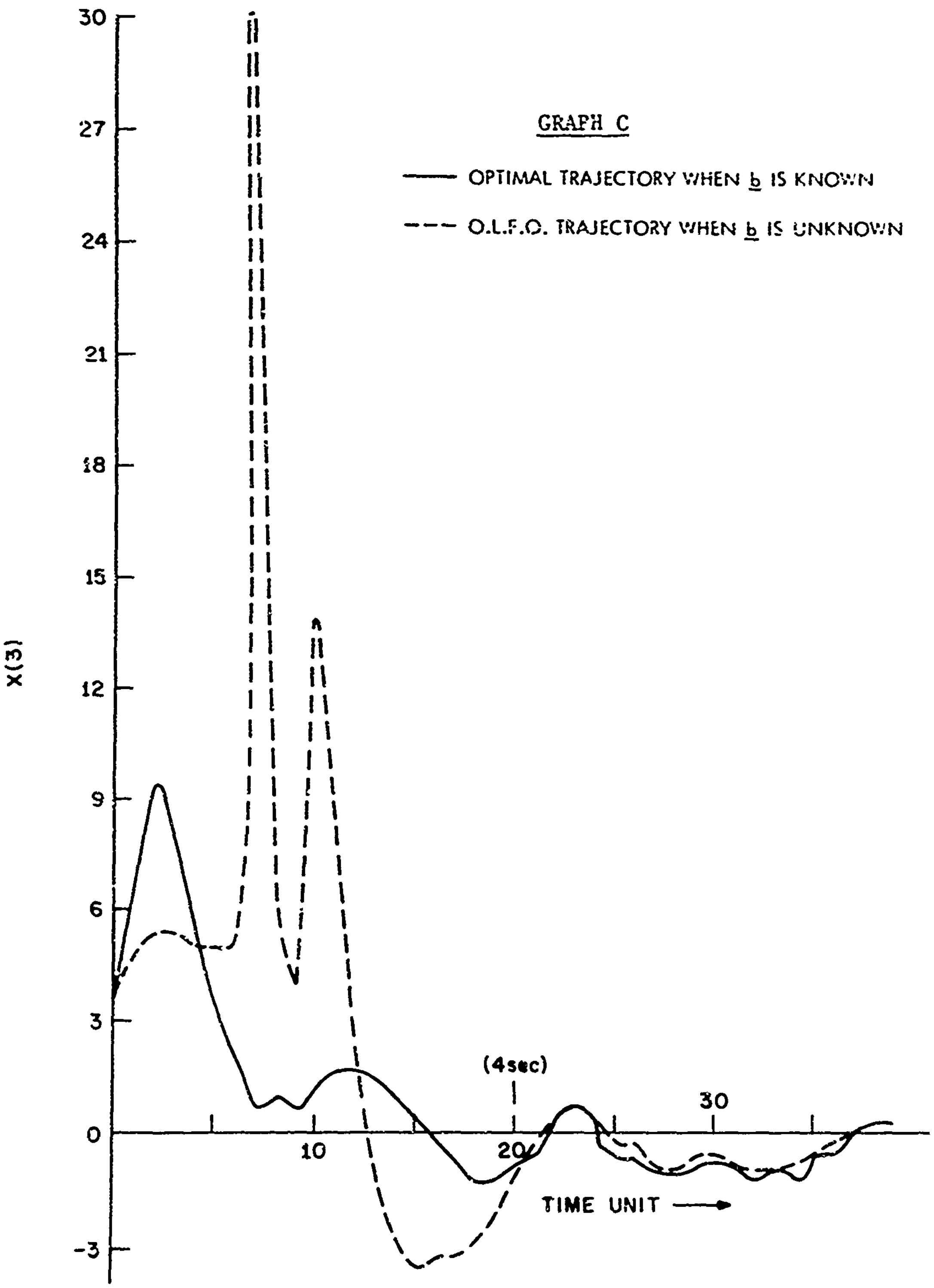


Fig. 7.2 (Continued)

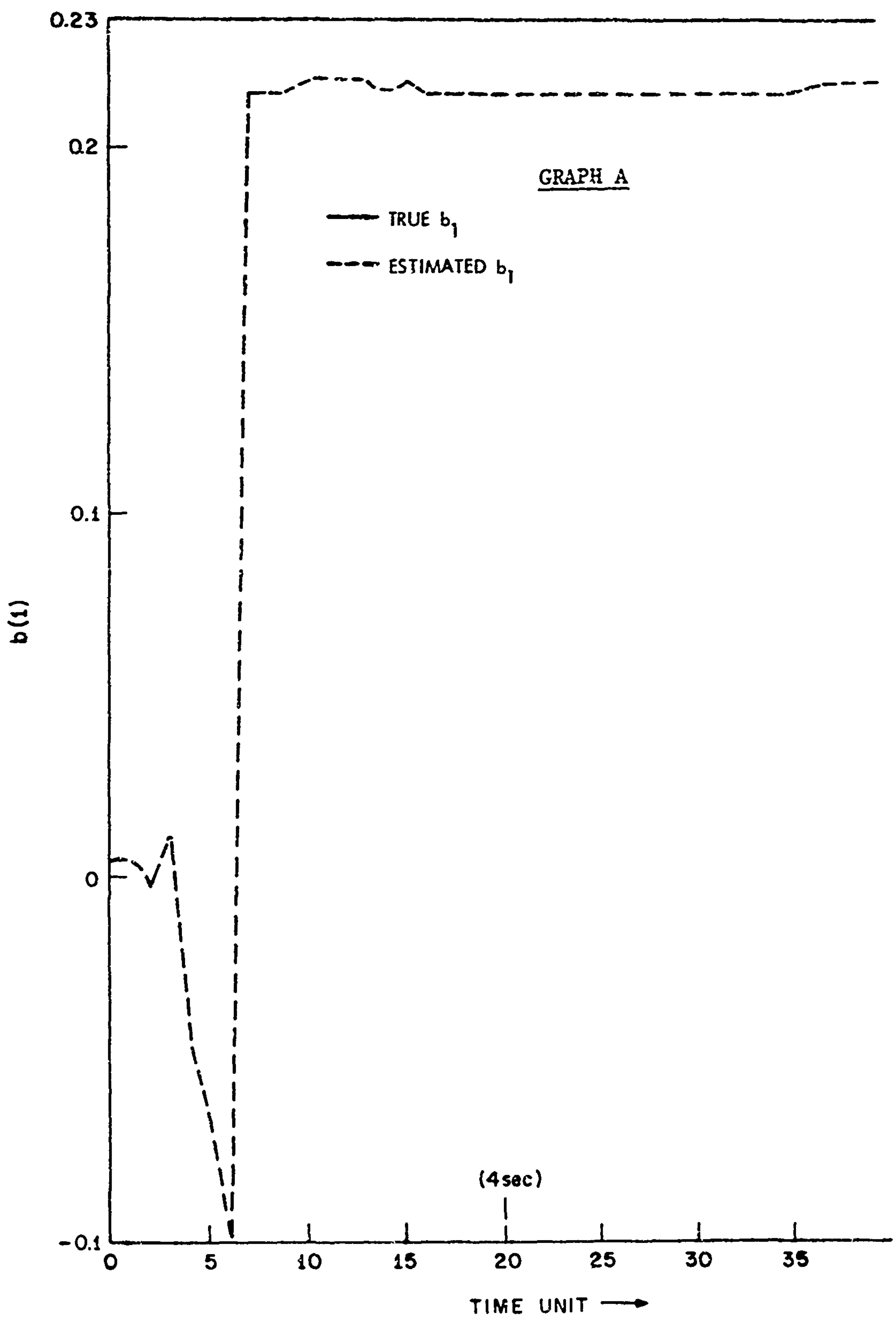


Fig. 7.3 ESTIMATE OF THE UNKNOWN GAIN VECTOR. THE SYSTEM BEING CONSIDERED HAS SYSTEM FUNCTION $\frac{(S+3)\ (S+2)}{(S-1)\ (S^2+2S+5)}$

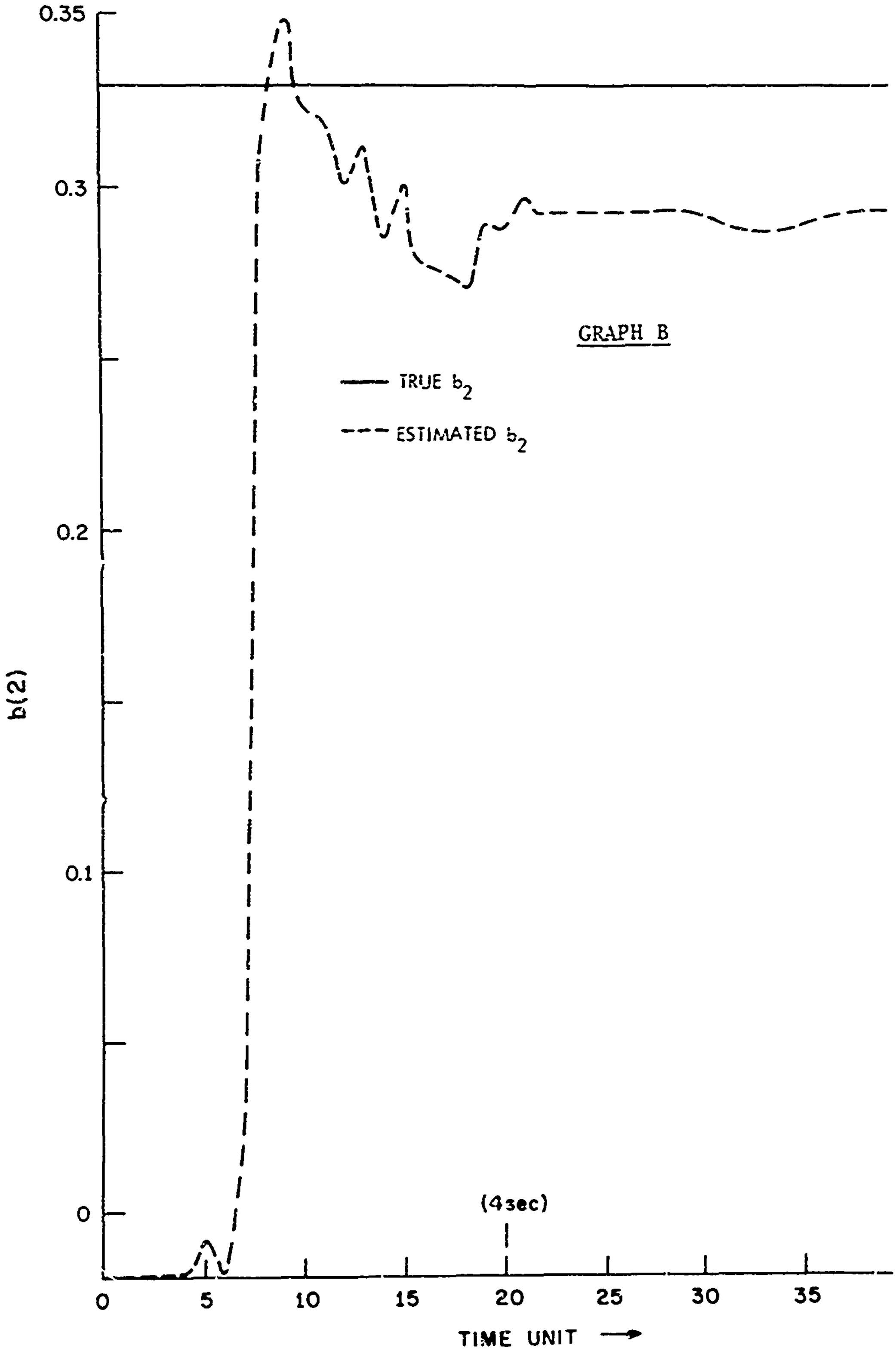


Fig. 7.3 (Continued)

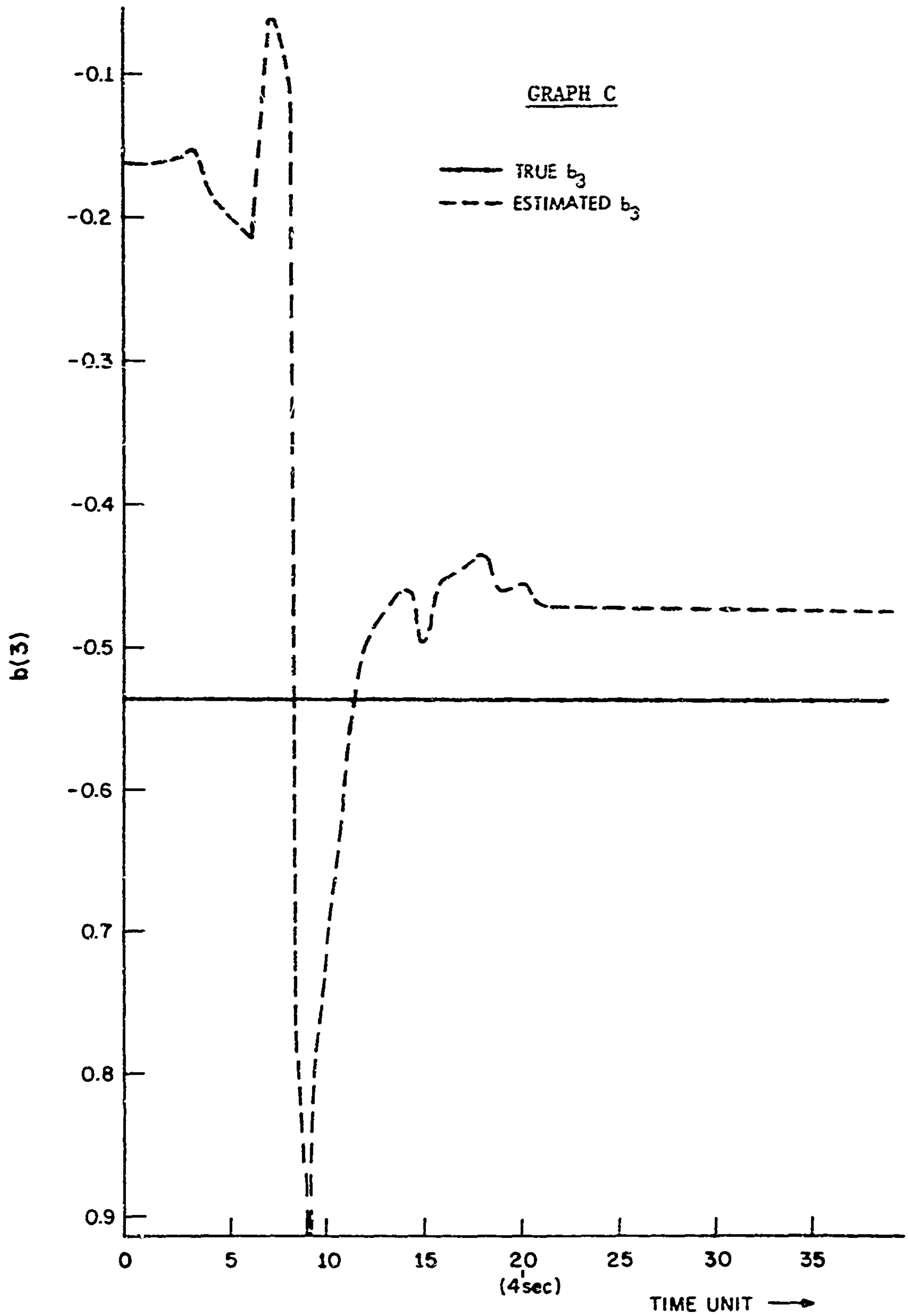


Fig. 7.3 (Continued)

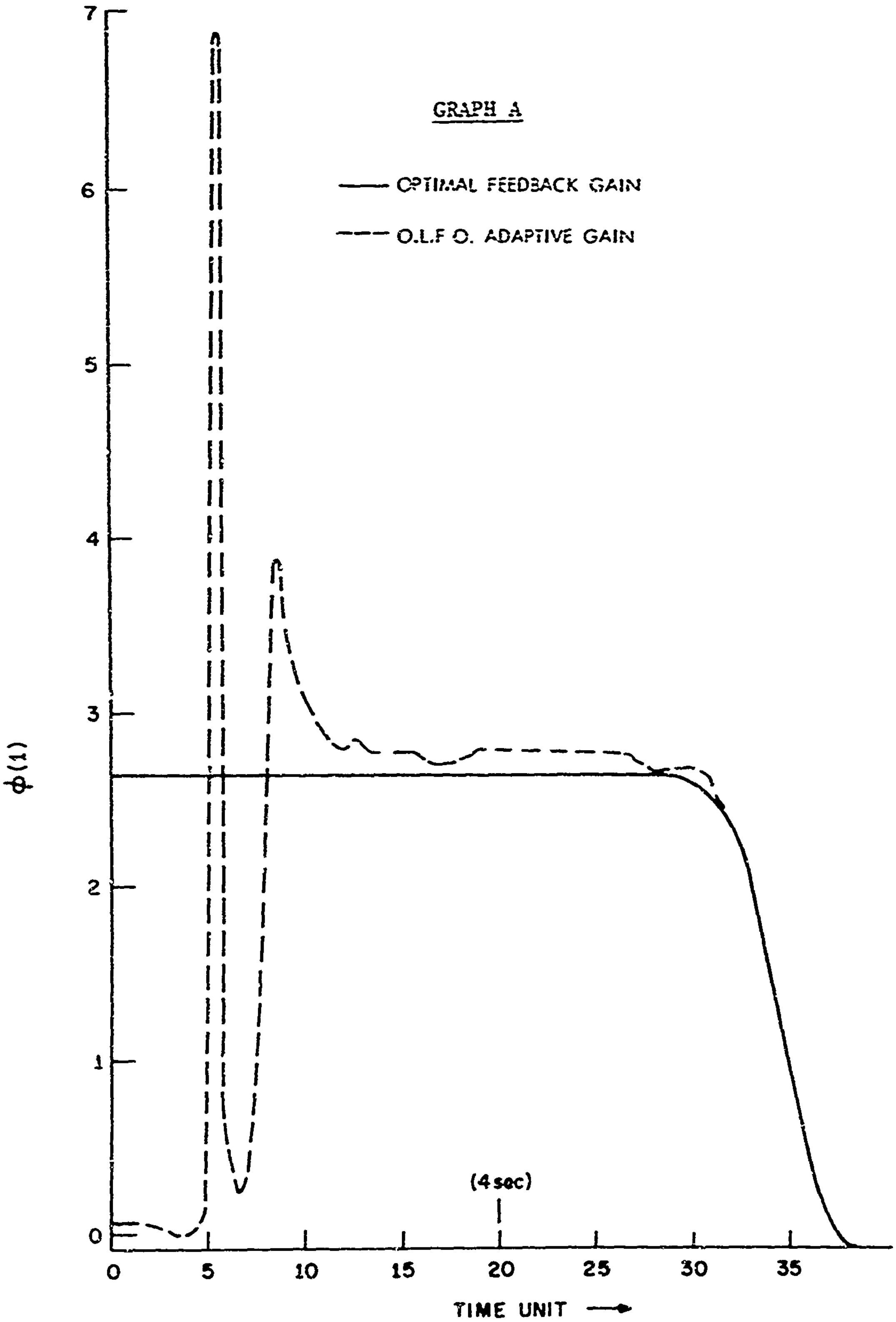


Fig. 7.4 COMPARISON BETWEEN THE OPTIMAL FEEDBACK GAINS AND THE ADAPTIVE O.L.F.O. GAINS. THE SYSTEM BEING CONSIDERED HAS SYSTEM FUNCTION $\frac{(S+3)\ (S+2)}{(S-1)\ (S^2+2S+5)}$

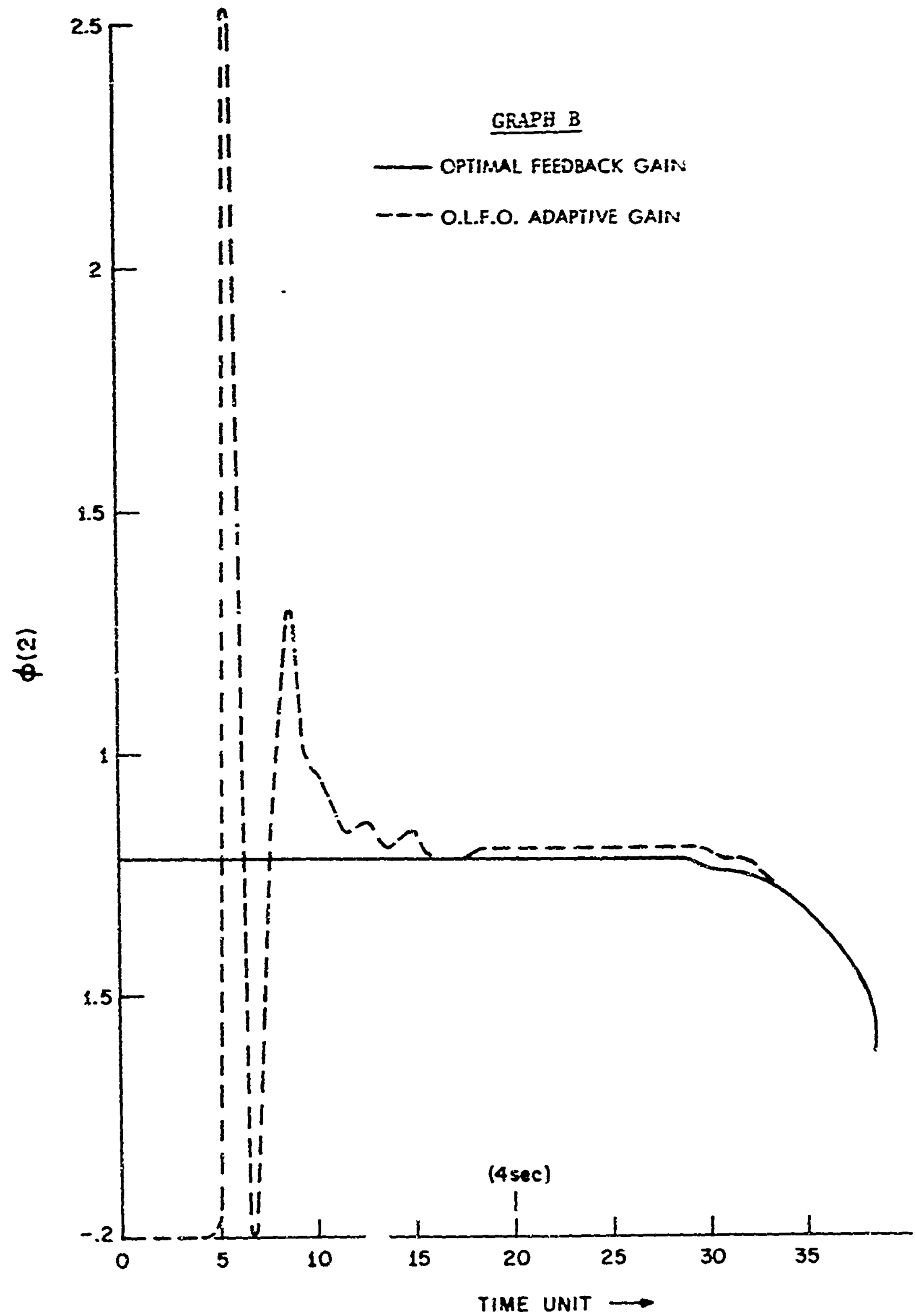


Fig. 7.4 (Continued)

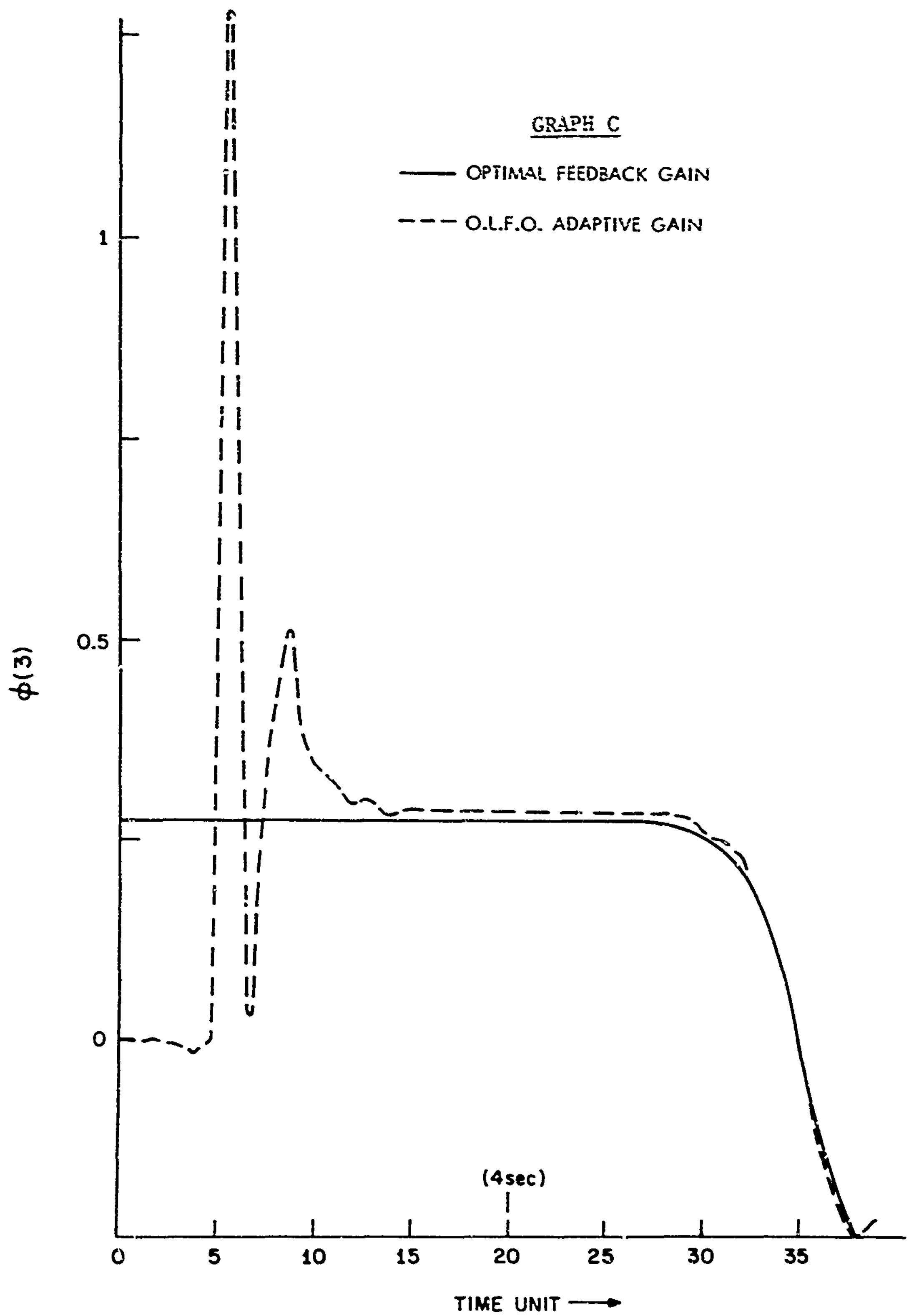


Fig. 7.4 (Continued)

i.e. we started out with an initial guess that the system has no zeroes. The final time is $N = 40$.

Many computer runs have been made on the same system with different noise samples. The plots for one particular sample experiment, which represents a fairly good average behavior, are shown in Figs. 7.2-7.4. From the experimental data (which are not shown completely), we can obtain a rough idea about the behavior of the suboptimal O.L.F.O. control system.

From the experiments, it was found that in the beginning, the O.L.F.O. adaptive gain is approximately zero (Fig. 7.4) and the O.L.F.O. trajectory follows closely to the input-free trajectory (Fig. 7.2). The diverging phenomenon is detected by the identifier; controls of considerably high magnitude are then applied for a few steps. This is indicated by the fact that there are sharp jumps in the state trajectories. Experiments show that these jumps are not caused by bad noise sample because the same phenomenon appears in different sample runs at approximately the same time interval. The high magnitude control serves mainly for identification purposes, this is revealed by the fact that at the next time unit, the estimates of $\underline{b}$ closely agree with the true $\underline{b}$ (Fig. 7.3). As was predicted in chapter 6, section 6.3, the O.L.F.O. adaptive gains do converge to the truly optimum gains (Fig. 7.3). The correction term vs. time is not shown in the figure, but simulation results indicate that the correction term goes to zero very rapidly after the identification of $\underline{b}$ is essentially completed.

Another set of simulation experiments was carried out where we kept the same sample noise but varied the weighting h, $(h > 0)$. It was found from the experiments (not reported in here) that the maximum magnitude of the overshoot in the O.L.F.O. trajectories varied inversely with the value

of h; if h was large, we have relatively "lower" overshoots; whereas, if h was small, we had relatively high overshoots. Also, the experiments seem to indicate that the convergence rate and the final estimation error in $\underline{b}$ seem to depend on the value of h we chose; with large h, we have relatively slow convergence rate and relatively big final estimation error in $\underline{b}$; if h is small, we have a relatively fast convergence rate and relatively small final estimation error in $\underline{b}$.

In the next set of experiments, we kept the weighting fixed ($h = 0.1$), and repeated the first set of experiments with larger driving noise covariance ($r = 0.45$) while using the same observation noise sample. The experimental results (not reported in here) seem to indicate that the increase in driving noise covariance has little effect on the convergence rate of the O.L.F.O. control system.

It is of interest to find out whether the initial guess on b_f will be sensitive to the resulting O.L.F.O. control system. We carried out a set of experiments where we fixed

$$\underline{b}_f = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \quad ; \quad \underline{A}_f = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 5 & -3 & -1 \end{bmatrix} \tag{7.17}$$

The transfer function is

$$H_1'(s) = \frac{-1}{(s-1)(s^2 + 2s + s)} \tag{7.18}$$

The initial condition on $\underline{x}_f(0)$ was kept fixed, and using the same sample noise, we varied our initial guess in $\underline{b}_f$. The same runs seem to indicate that though the sample O.L.F.O. trajectory varied with different initial guesses in b_f; the convergence rate was quite insensitive to the guess in $\underline{b}_f$.

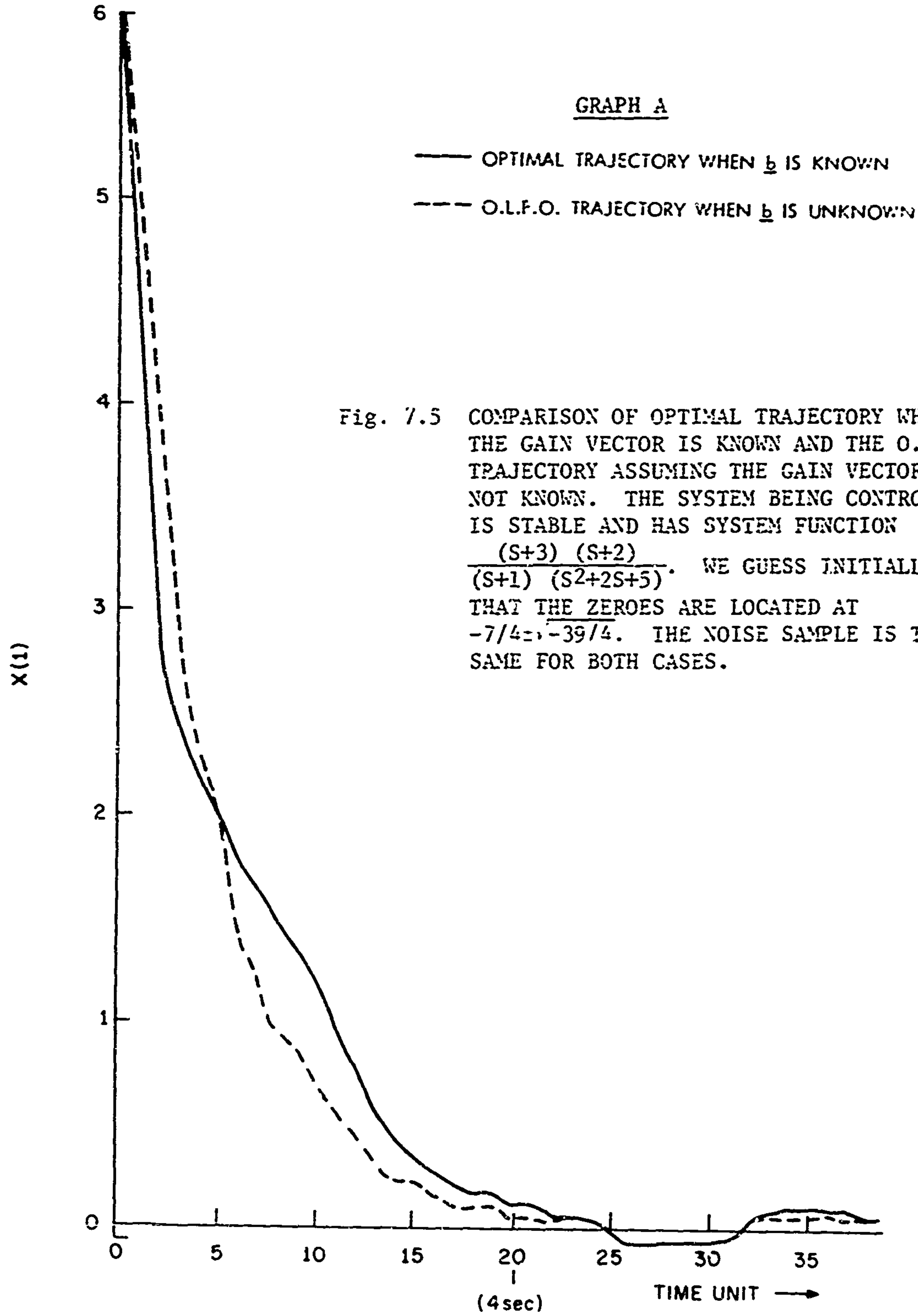


Fig. 7.5 COMPARISON OF OPTIMAL TRAJECTORY WHEN THE GAIN VECTOR IS KNOWN AND THE O.L.F.O. TRAJECTORY ASSUMING THE GAIN VECTOR IS NOT KNOWN. THE SYSTEM BEING CONTROLLED IS STABLE AND HAS SYSTEM FUNCTION $\frac{(S+3)\ (S+2)}{(S+1)\ (S^2+2S+5)}$. WE GUESS INITIALLY THAT THE ZEROES ARE LOCATED AT $-7/4 \pm \sqrt{-39}/4$. THE NOISE SAMPLE IS THE SAME FOR BOTH CASES.

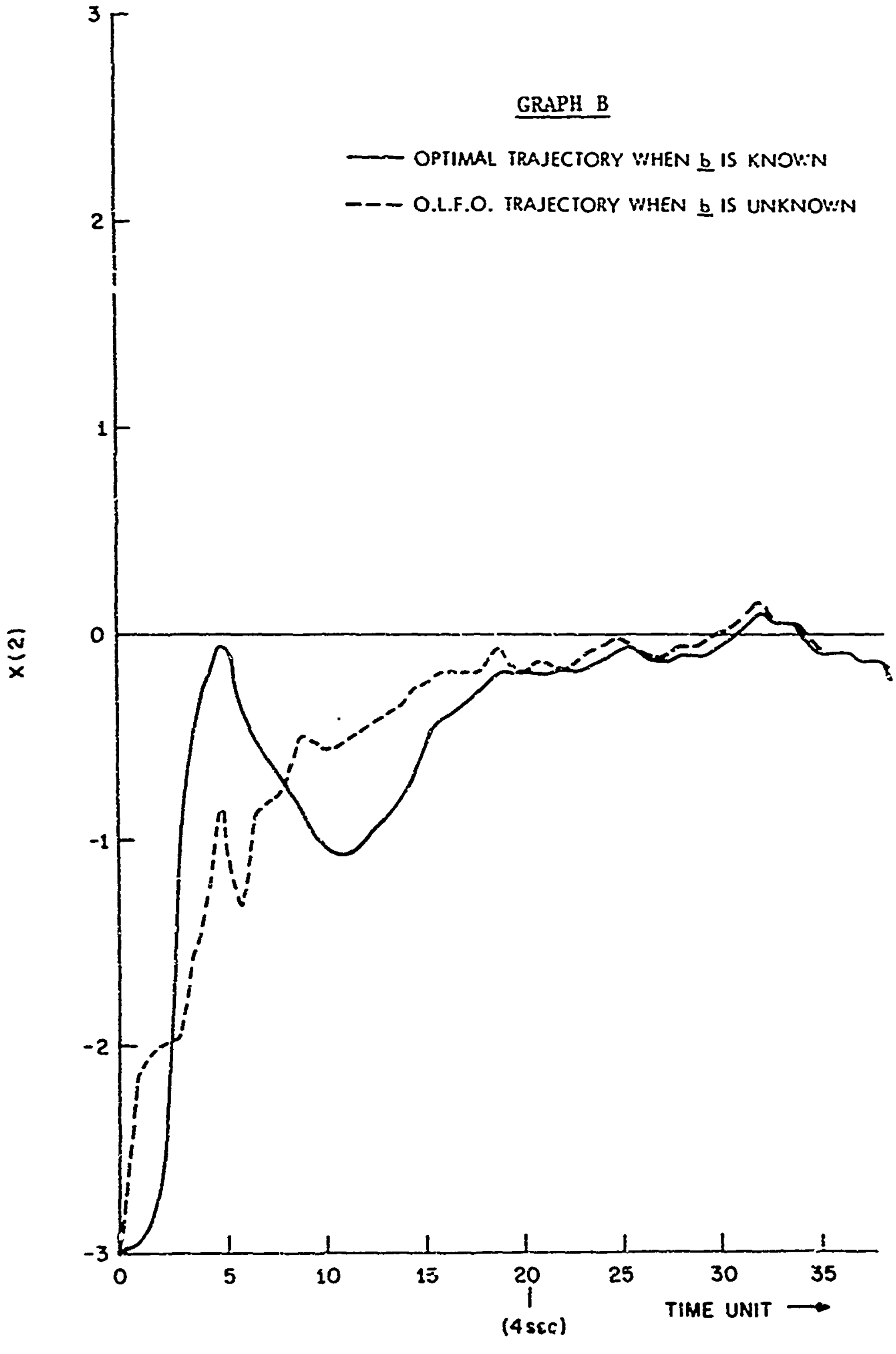


Fig. 7.5 (Continued)

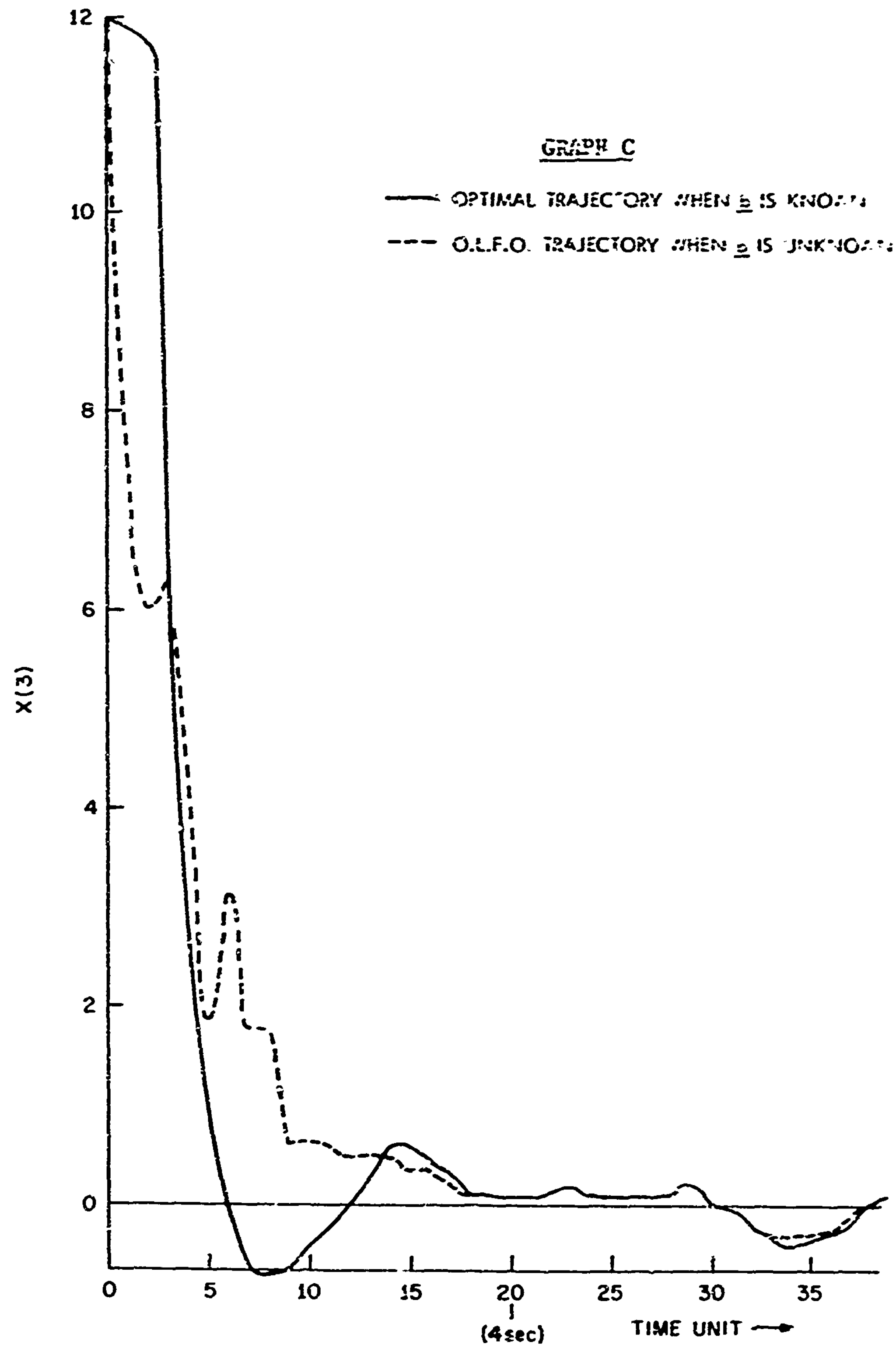


Fig 7.5 (Continued)

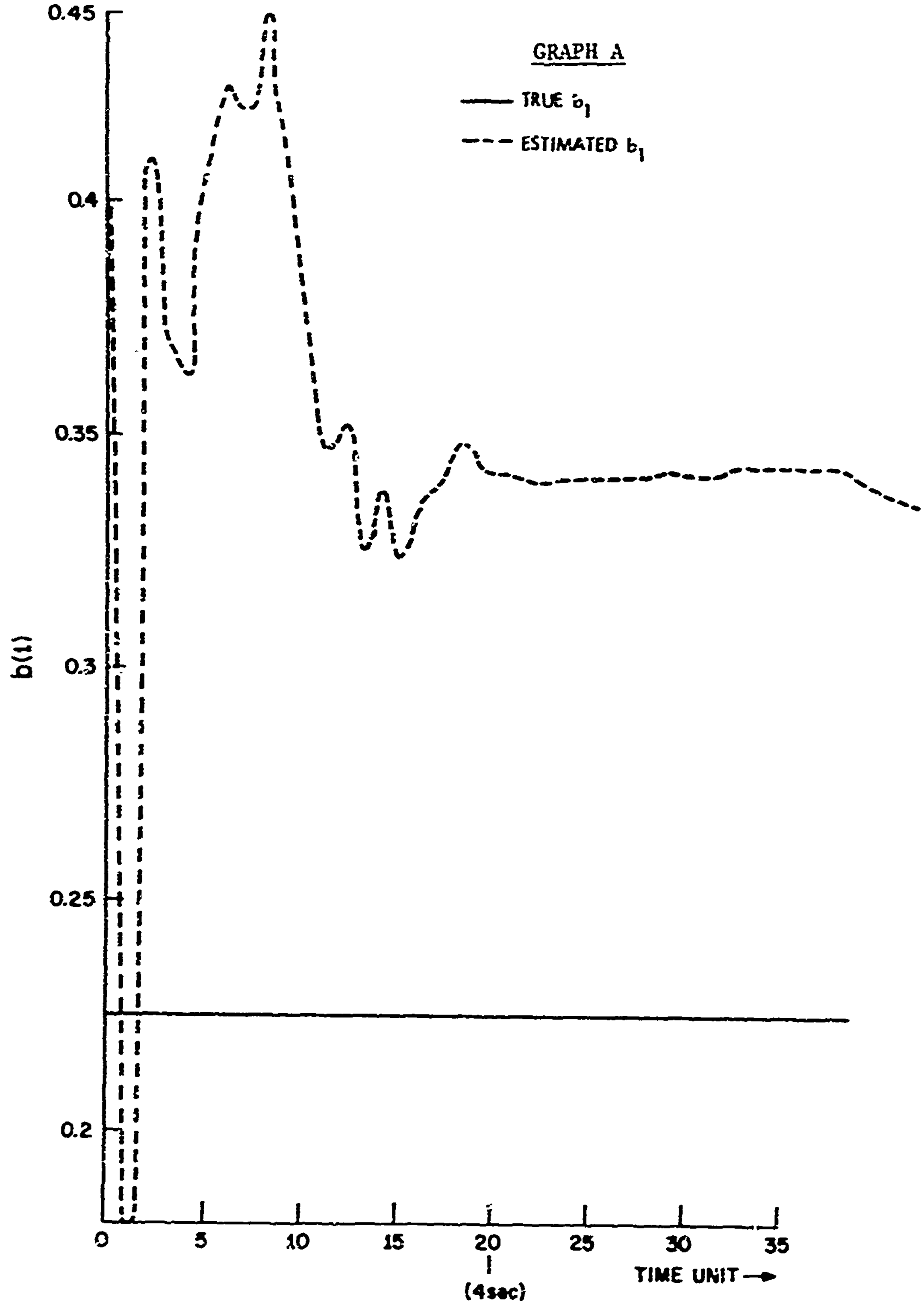


Fig. 7.6 ESTIMATE OF GAIN VECTOR. THE SYSTEM BEING CONSIDERED HAS SYSTEM FUNCTION $\frac{(S+3)\ (S+2)}{(S+1)\ (S^2+2S+5)}$. WE GUESS INITIALLY THAT THE ZEROES ARE LOCATED AT $-7/4 \pm \sqrt{-39/4}$.

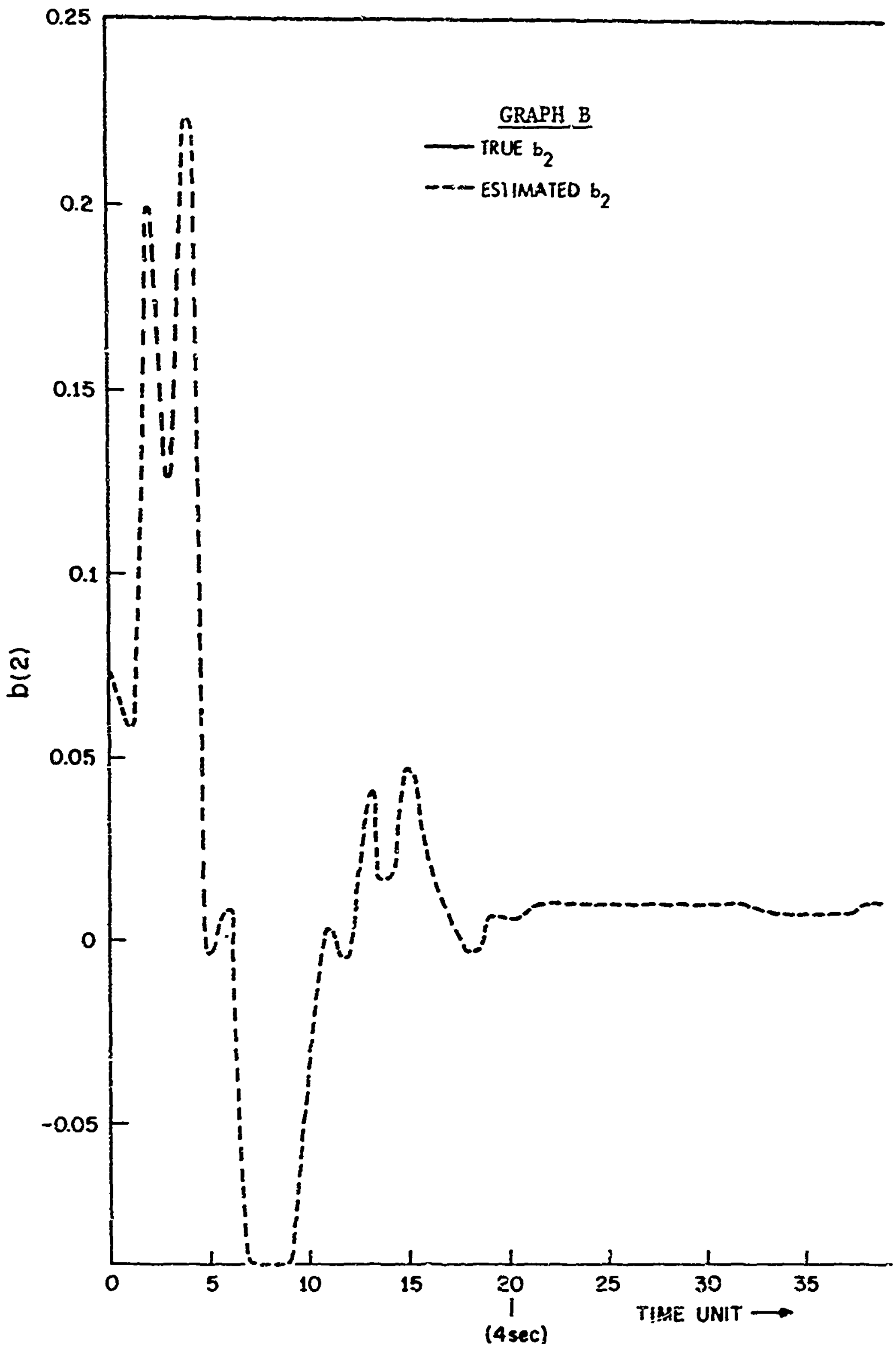


Fig. 7.6 (Continued)

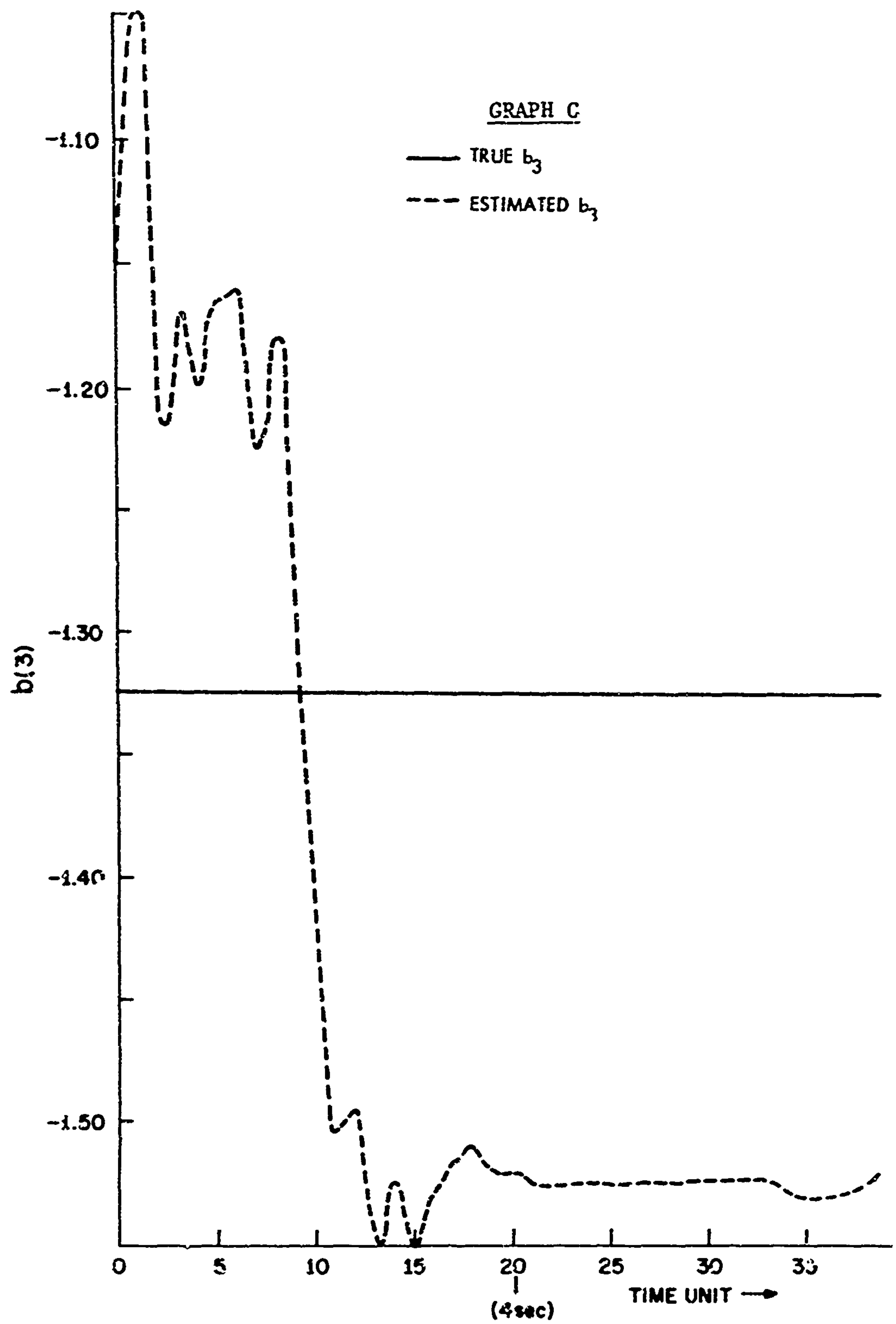


Fig. 7.6 (Continued)

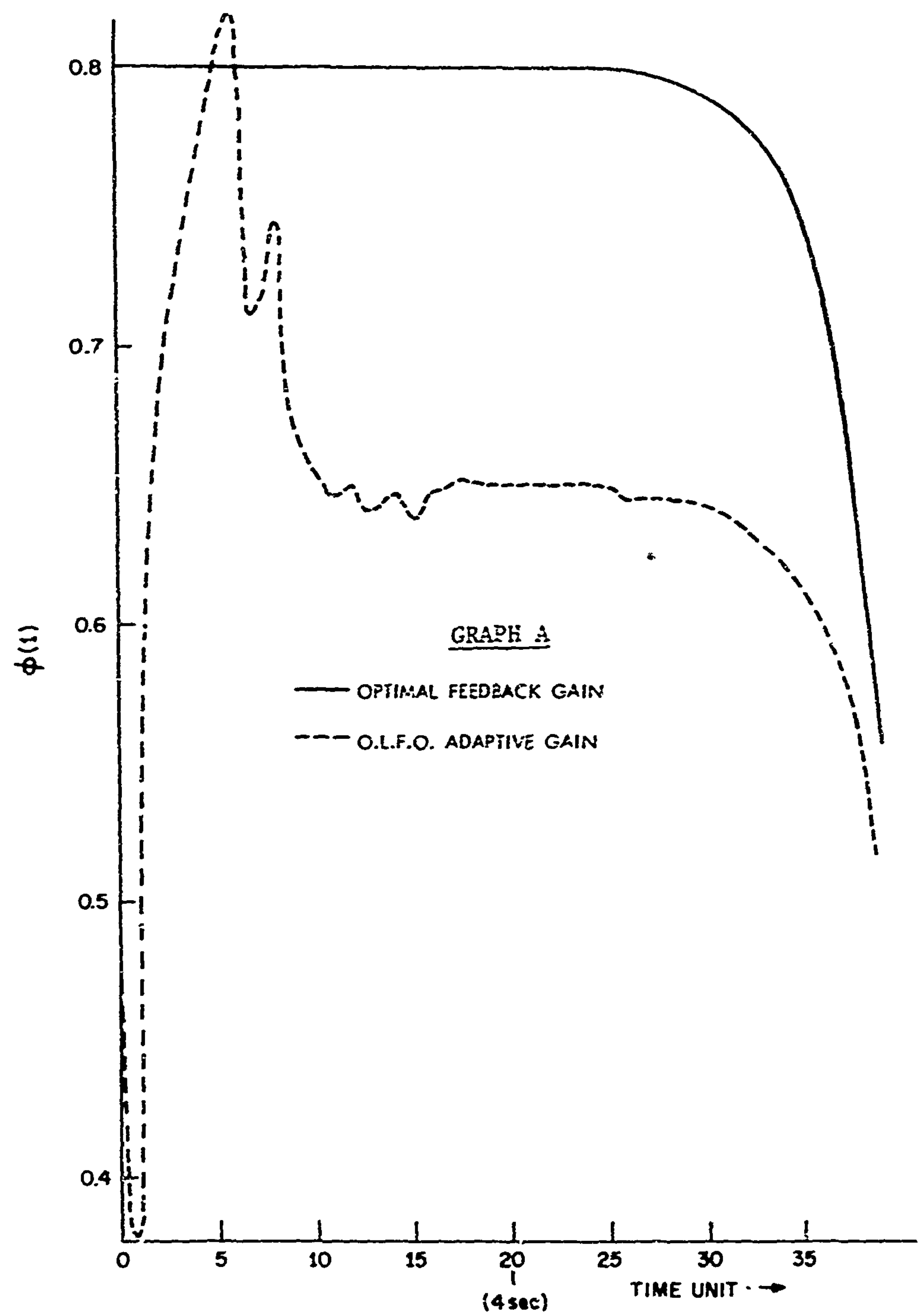


Fig. 7.7 COMPARISON BETWEEN OPTIMAL FEEDBACK GAIN AND O.L.F.O. ADAPTIVE GAIN. THE SYSTEM BEING CONSIDERED HAS SYSTEM FUNCTION $\frac{(S+3)(S+2)}{(S+1)(S^2+2S+5)}$. WE GUESS INITIALLY THAT THE ZEROES ARE LOCATED AT -7/4, -39/4

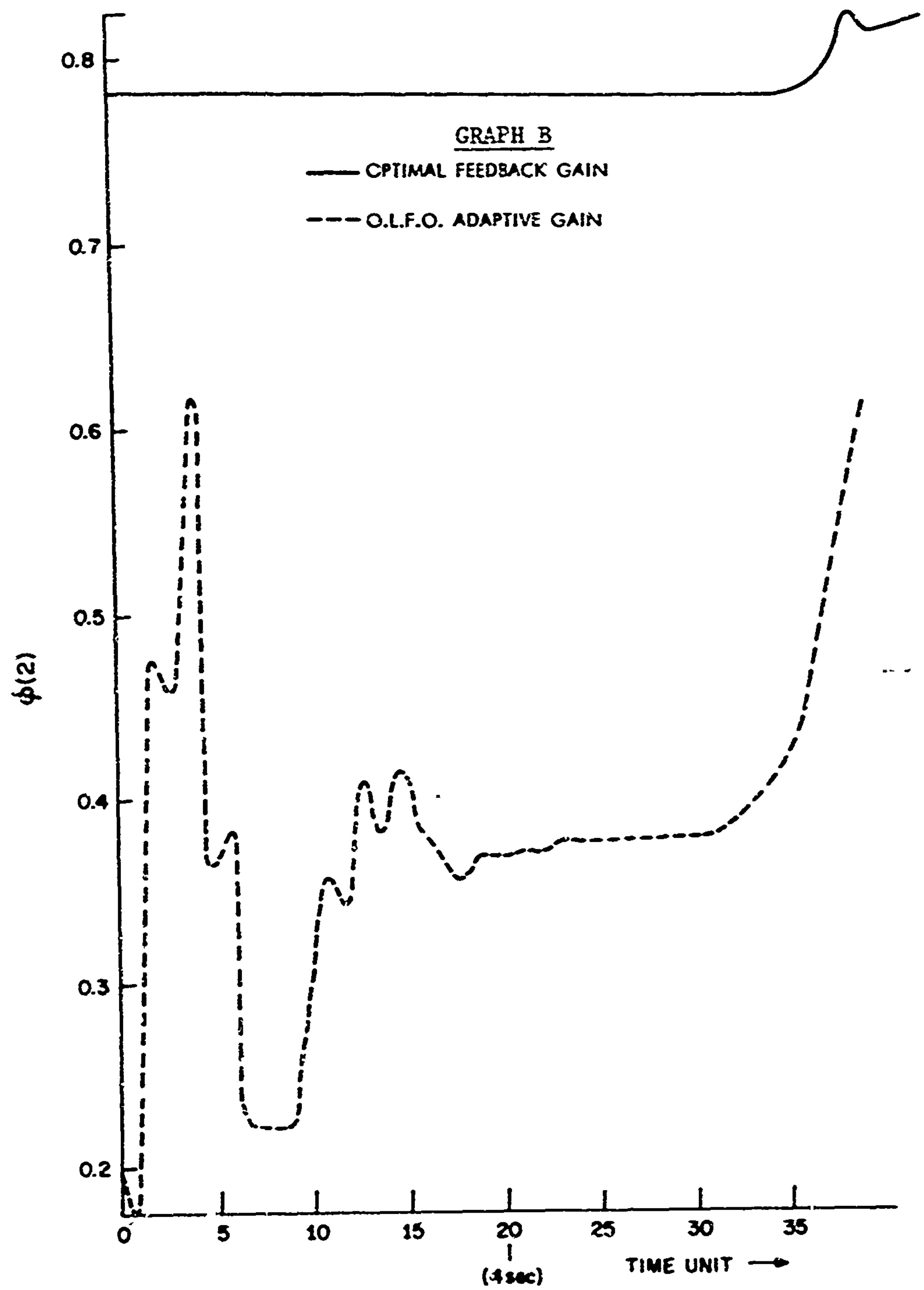


Fig. 7.7 (Continued)

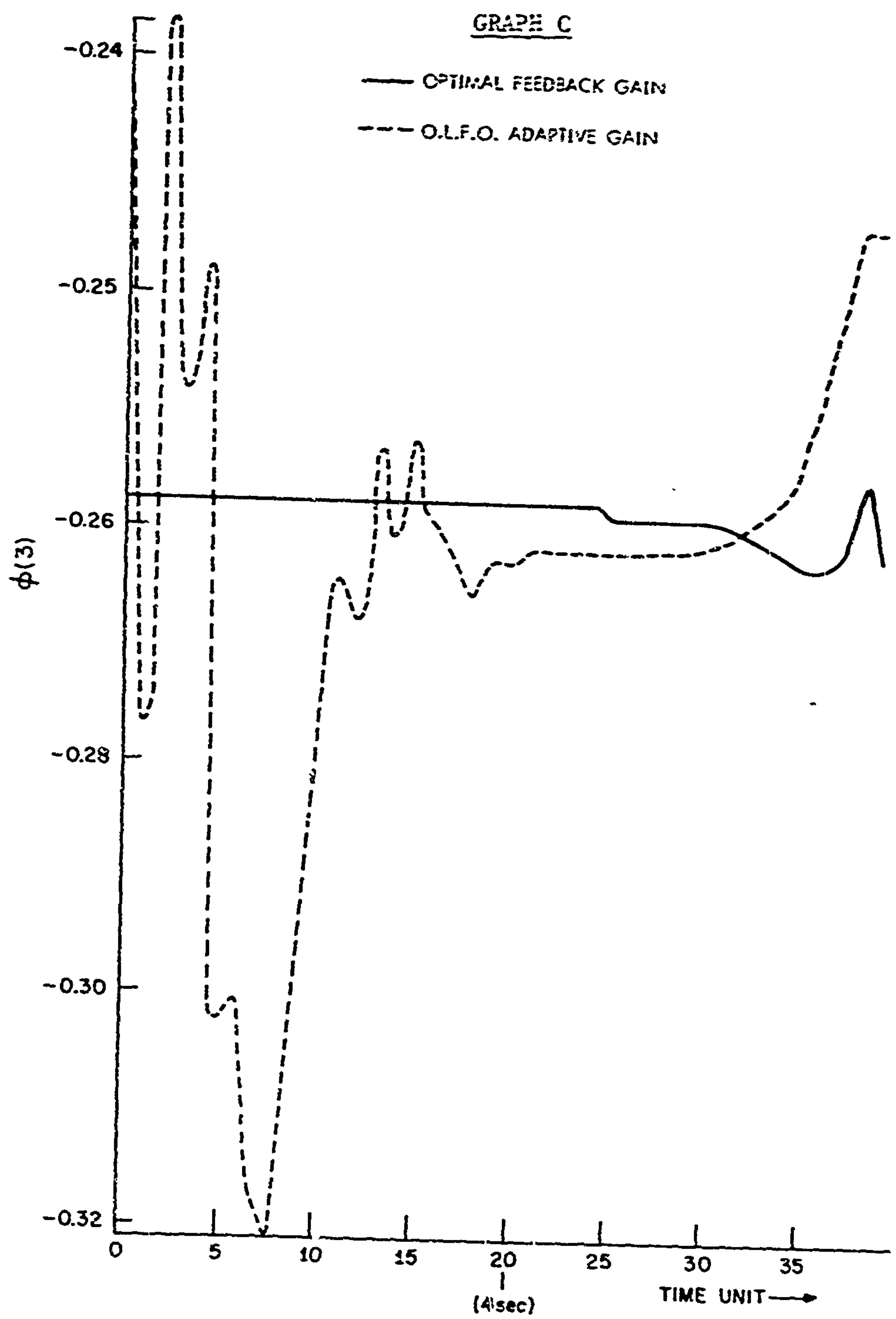


Fig. 7.7 (Continued)

Example 2: Stable System

It is assumed that

$$\underline{A}_f = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -5 & -7 & -3 \end{bmatrix} \quad ; \quad \underline{b}_f = \begin{bmatrix} 1 \\ 2 \\ -7 \end{bmatrix} \quad ; \quad \underline{x}_f(0) = \begin{bmatrix} 6 \\ -3 \\ 12 \end{bmatrix} \tag{7.19}$$

The true transfer function for the system is (Fig. 7.1)

$$H_2(s) = \frac{(s + 3)(s + 2)}{(s + 1)(s^2 + 2s + s)} \tag{7.20}$$

The system is stable.

In the first set of experiments, we initially guess

$$\hat{\underline{b}}_f(0/0) = \begin{bmatrix} 2 \\ 1 \\ -6 \end{bmatrix} \tag{7.21}$$

i.e. that the zeroes are located at $-\frac{7}{4} + \sqrt{\frac{-39}{4}}$ and $-\frac{7}{4} - \sqrt{\frac{-39}{4}}$. The weighting on the control is $h = 1$. We take the final time $N = 40$.

Sample runs for the same system with same initial guess (7.21) were made and the plots for one particular sample are shown in Figs. 7.5-7.7. As opposed to the unstable case, the O. L. F. O. adaptive gain is some nonzero vector, and so the value of the O.L.F.O. control is not zero at the beginning (Fig. 7.7). The control is used both for identification and control purposes. The system is stable, and since no large magnitude control is applied, the O.L.F.O. trajectory decays down to zero (see Fig. 7.5). This decaying phenomenon is noticed by the identifier, and thus the control is kept near zero to save energy. Therefore, after a certain time interval, when the O.L.F.O. trajectory goes near the origin, the O.L.F.O. control will remain zero for most of the time. The system behaves almost like an input-free system. In fact, this is also what the

truly optimum system will do. We note from Fig. 7.6 that the identification process of the unknown gain $\underline{b}$ stops at about $k = 20$, which is the approximate time unit when the O.L.F.O. state trajectory begins to stay around zero. If we consider control over an infinite interval (say using a window-shifting approach) we may expect awfully slow convergence rate in the estimation of $\underline{b}$ to the true $\underline{b}$, and a slow convergence rate of O.L.F.O. control system to truly optimum control system.

In the second set of experiments, we have the same noise samples as before but starting with the initial condition

$$\underline{x}_f(0) = \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix} \tag{7.22}$$

The initial guess on $\underline{b}_f$ was

$$\hat{\underline{b}}_f(0/0) \quad \begin{bmatrix} 0 \\ 0 \\ -2 \end{bmatrix} \tag{7.23}$$

i.e. there are no zeroes. The weighting on the control is $h = 1$, and we take the final time $N = 60$. The plots for one typical sample experiment are shown in Figs. 7.8-7.10. (The sample noise for the sample run shown in Figs. 7.8-7.10 is the same as that shown in Figs. 7.5-7.7.) Comparing this set of experiments with the last, we note that more or less the same phenomenon occurred in both sets of experiments. The final estimate in $\underline{b}$ is way off its true value, in fact $\hat{b}_1(k/k)$ and $\hat{b}_2(k/k)$ are opposite in sign with those of b_1 and b_2 respectively; but interestingly enough the adaptive gains are adjusted accordingly so that the values of the O.L.F.O. control sequence and the truly optimal control sequence are almost the same. This set of experiments indicates yet slower convergence (if there is any).

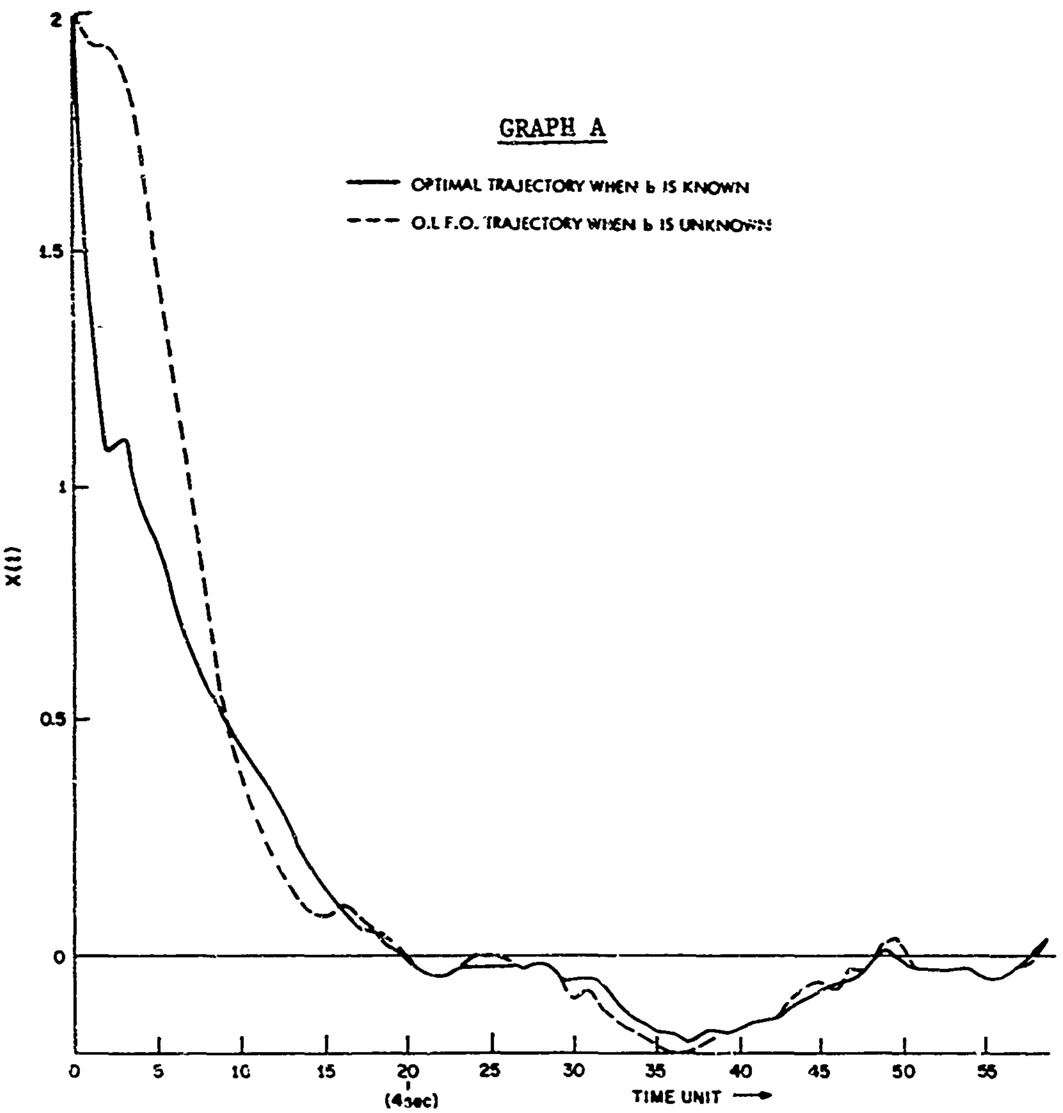


Fig. 7.8 COMPARISON BETWEEN OPTIMAL TRAJECTORY WHEN THE GAIN VECTOR IS KNOWN AND THE O.L.F.O. TRAJECTORY ASSUMING THE GAIN VECTOR IS UNKNOWN. THE SYSTEM BEING CONTROLLED HAS SYSTEM FUNCTION $\frac{(S+3)\ (S+2)}{(S+1)\ (S^2+2S+3)}$. WE GUESS INITIALLY THAT THERE ARE NO Z EROES. THE NOISE SAMPLE IS THE SAME FOR BOTH CASES.

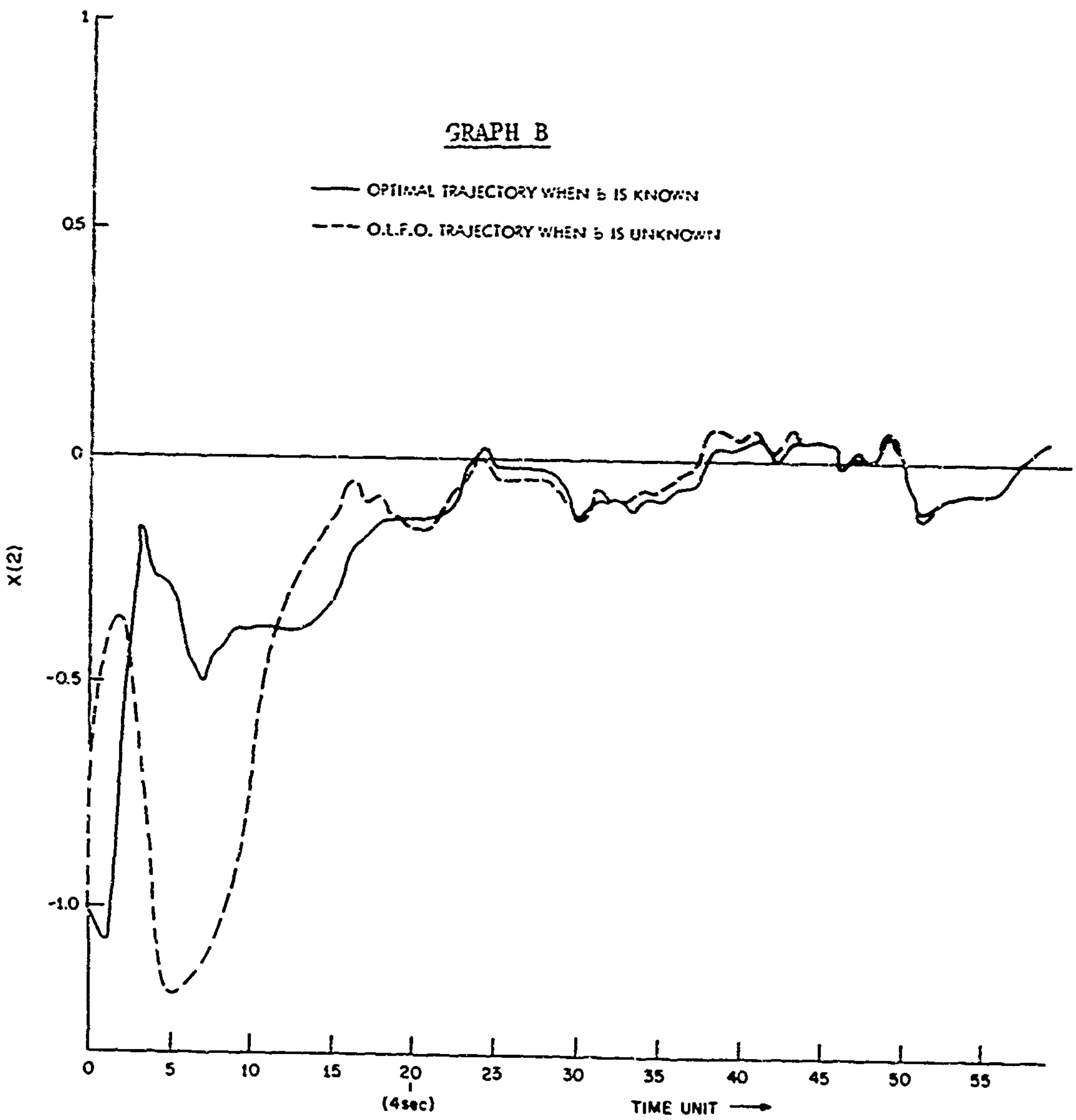


Fig. 7.8 (Continued)

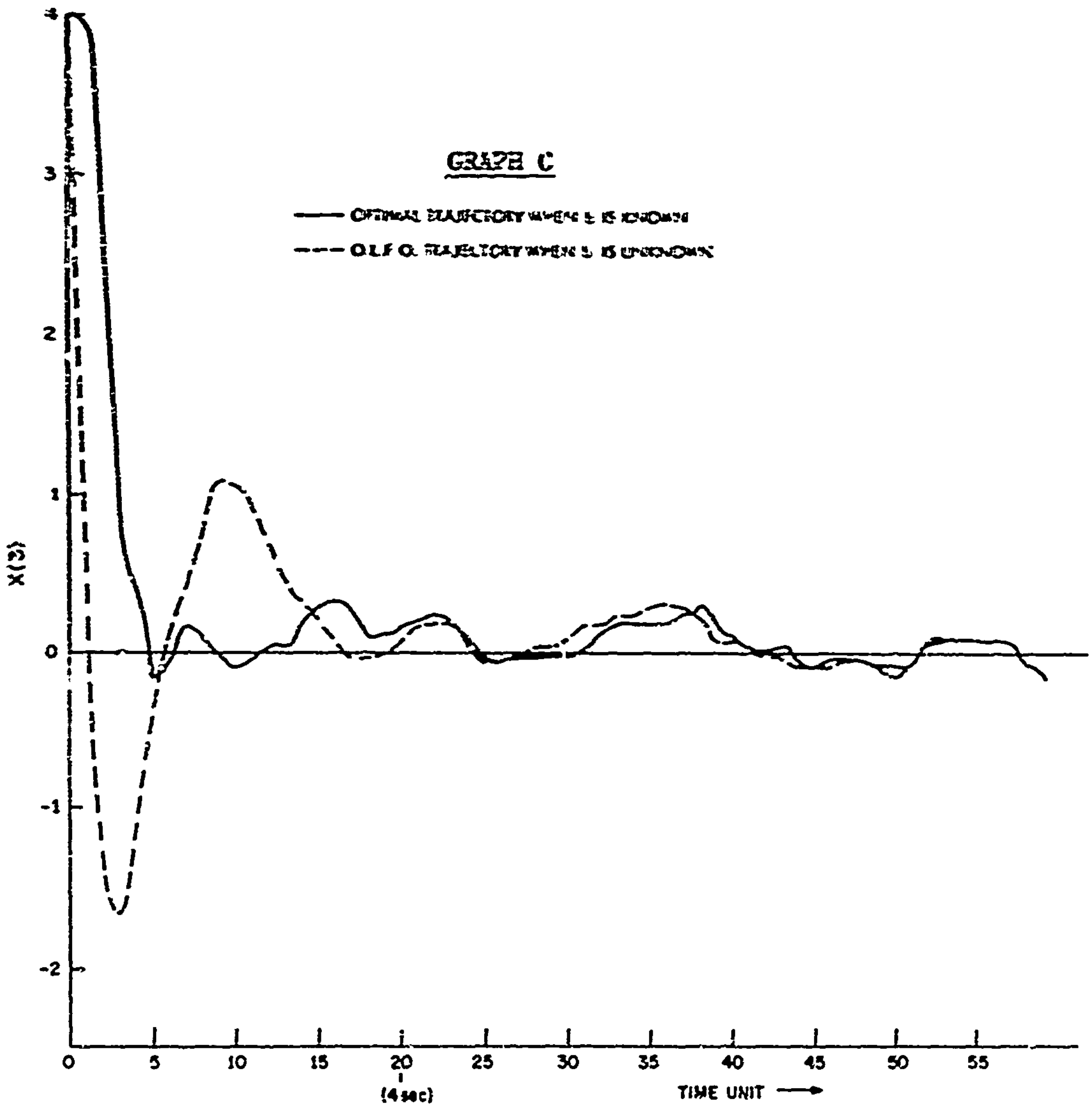


Fig. 7.8 (Continued)

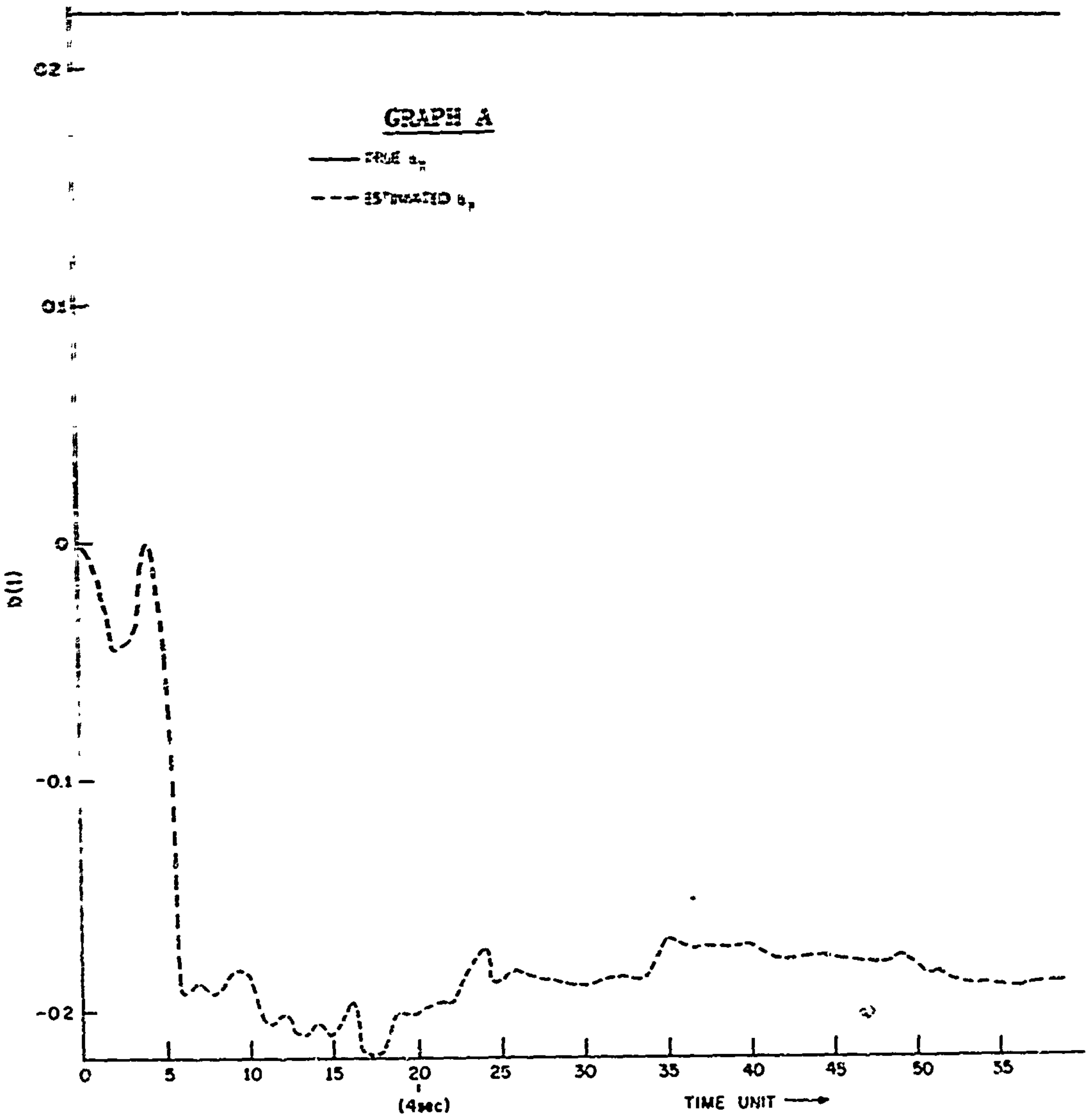


Fig. 7.9 ESTIMATE OF THE GAIN VECTOR. THE SYSTEM BEING CONSIDERED HAS SYSTEM FUNCTION $\frac{(S+3)\ (S+2)}{(S+1)\ (S^2+2S+3)}$. THE INITIAL GUESS IS THAT THERE ARE NO ZEROES.

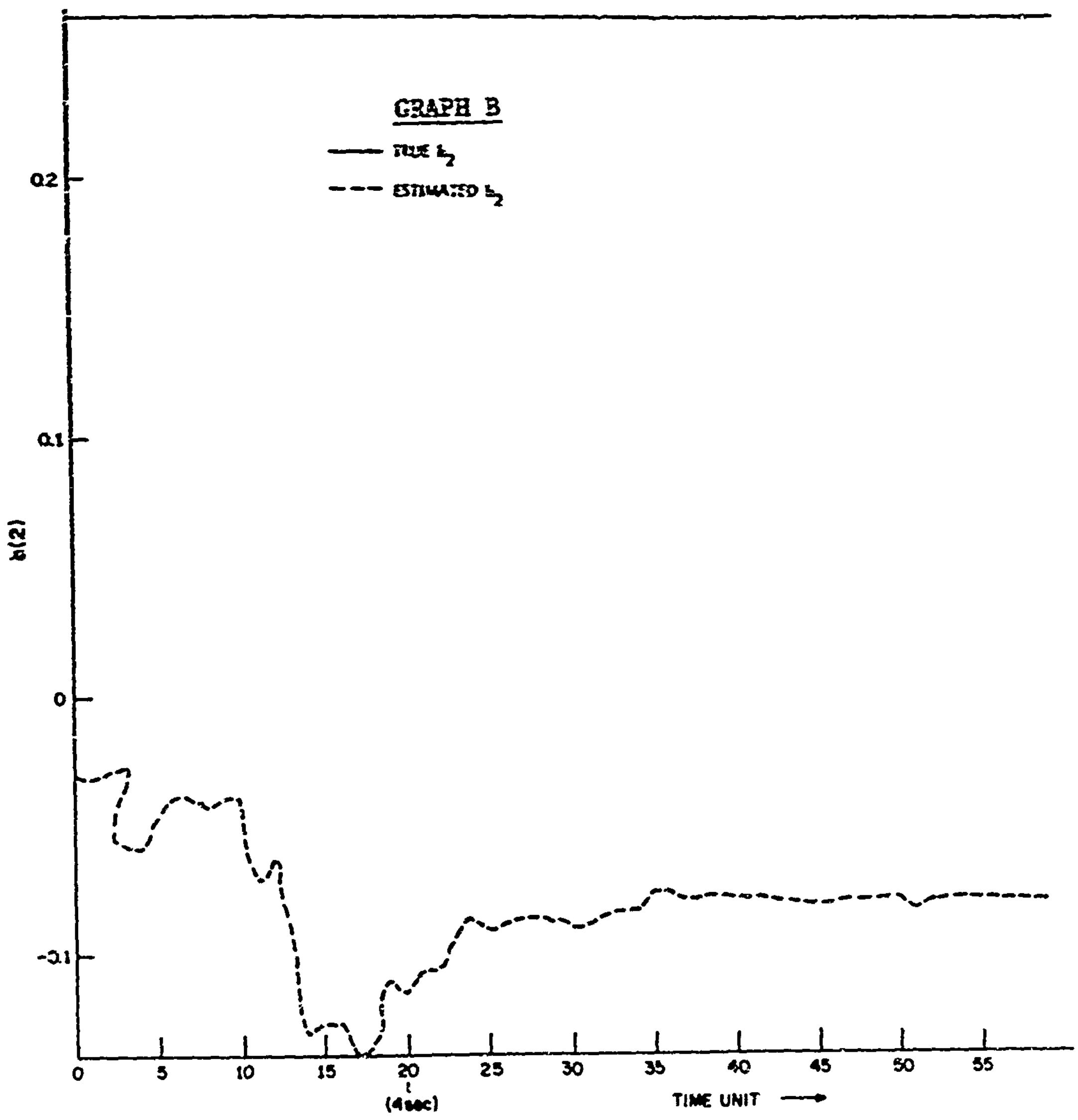
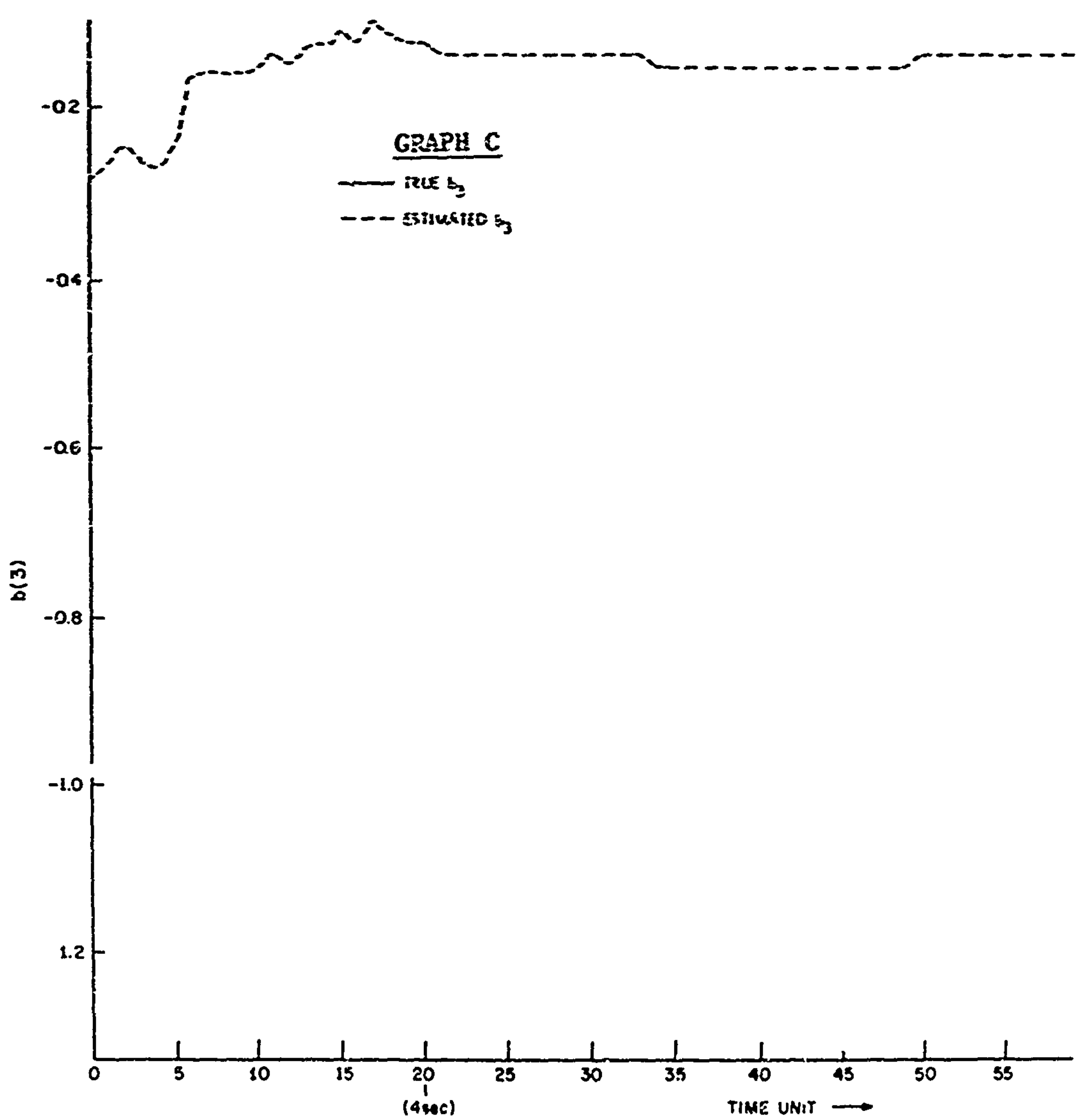


Fig. 7.9 (Continued)



ig. 7.9 (Continued)

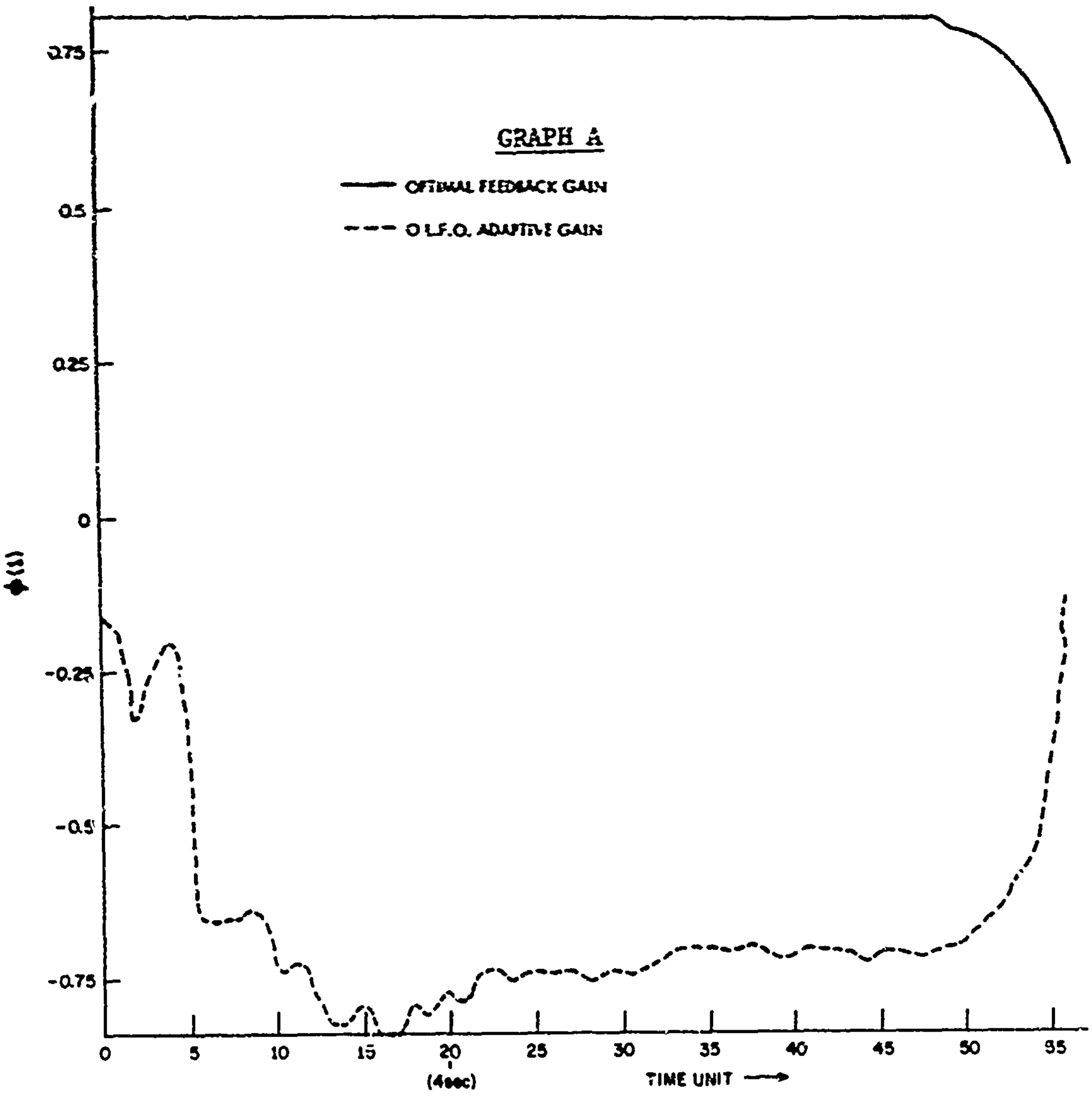


Fig. 7.10 COMPARISON BETWEEN OPTIMAL TO FEEDBACK GAIN AND O.L.F.O. ADAPTIVE GAIN. THE SYSTEM BEING CONSIDERED IS $\frac{(S+3)\ (S+2)}{(S+1)\ (S^2+2S+2)}$. THE INITIAL GUESS IS THAT THERE ARE NO ZEROES.

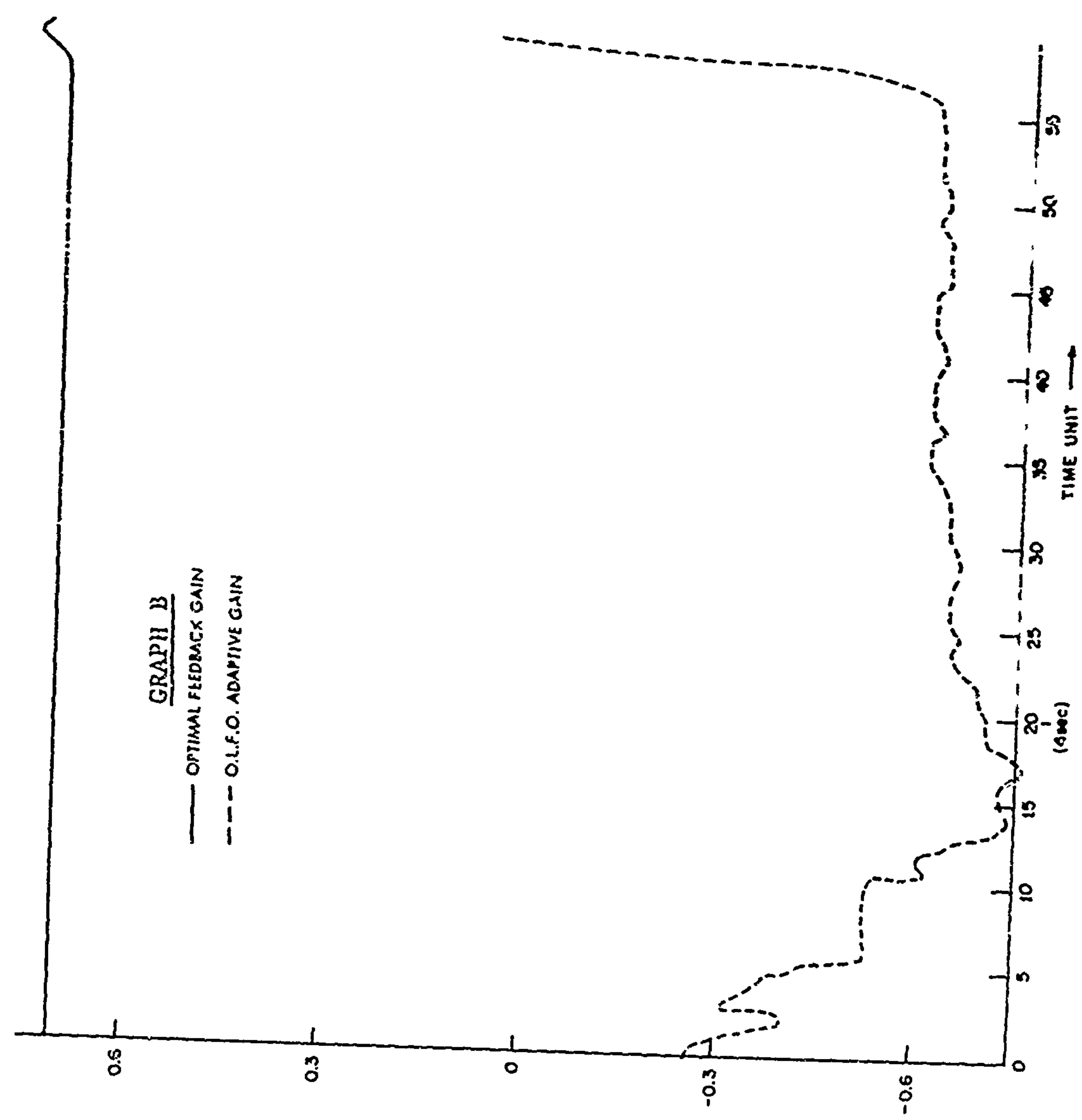


Fig. 7.10 (Continued)

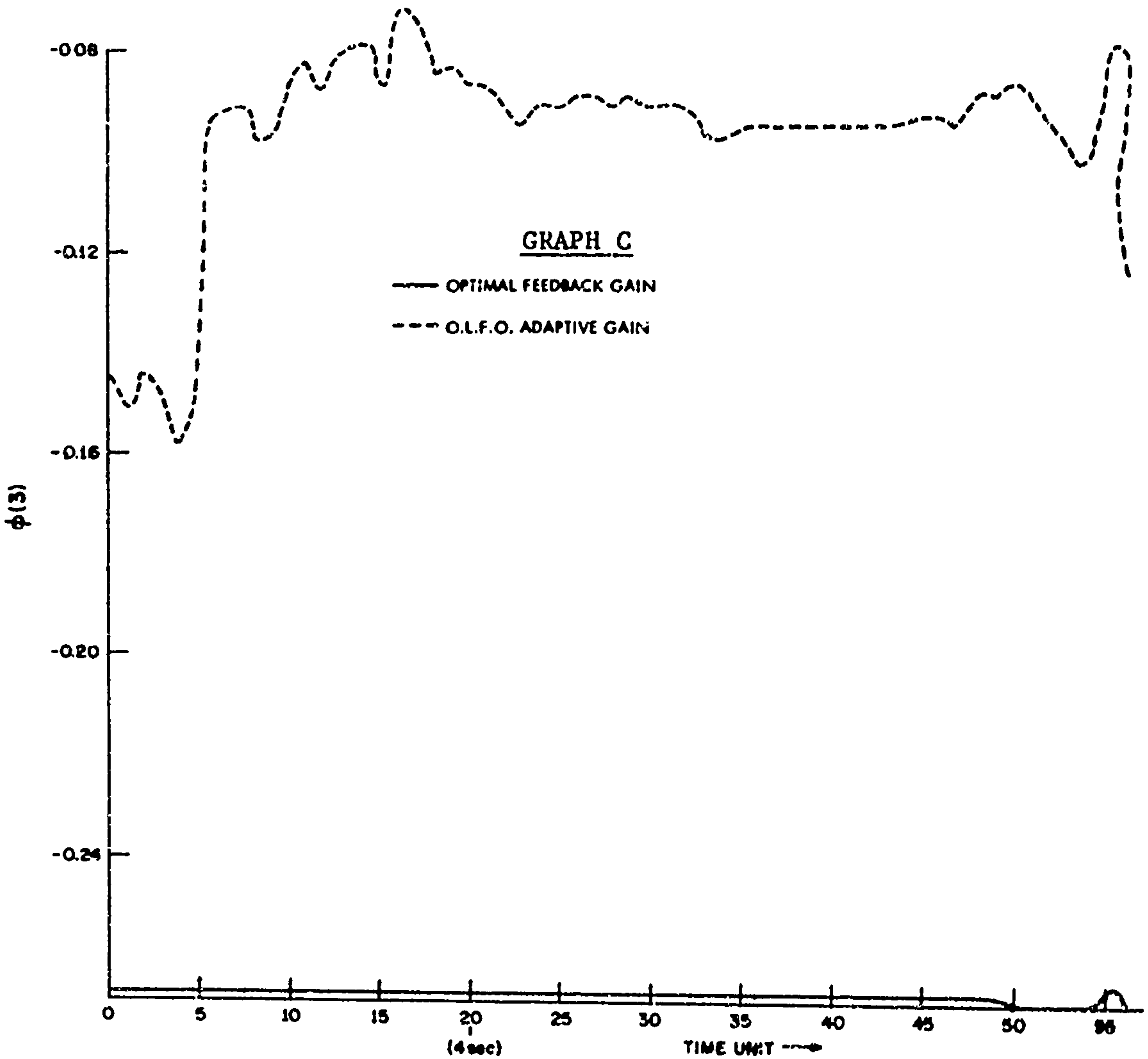


Fig. 7.10 (Continued)

Note that in both sets of experiments even if the estimate of $\underline{b}$ does not converge to the true $\underline{b}$, the truly optimal trajectory and O.L.F.O. trajectory are almost the same after the transient period.

Intuitively, the results are reasonable. Since we have not told the problem to identify $\underline{b}$, it will not do so unless the identification is absolutely necessary. The experimental results verified our theoretical deduction (see chapter 6, section 6.6).

The experiments seem to indicate that for stable system, the choice of initial guess will not greatly influence the O.L.F.O. trajectory, but will affect the convergence rate for the estimate in the gain parameters, $\underline{b}$.

Remark: In each set of experiments discussed above, the number of sample runs is not enough to enable us to draw specific statistical conclusions; yet the regularity in the sample runs enable us to draw some crude conclusions.

From the experiments, we may draw the following conclusions regarding the O.L.F.O. control system.

(1) The rate of convergence seems to be very dependent on the stability of the system. For unstable systems, the convergence rate seems to be faster compared to that for stable systems. This verifies our theoretical predictions made in chapter 6, section 6.6.

(2) It seems that large controls will help identification of the unknown gain parameters, and so convergence rate seems to relate directly to the magnitude of the control action. This again agrees with our intuitive remark made in chapter 6, section 6.3.

(3) For unstable systems, the rate of convergence seems to be fairly independent of the initial guess on the unknown gain,

whereas for stable systems, the convergence rate may be quite <u>dependent</u> on the initial guess on the unknown gain.

(4) For unstable systems, the O.L.F.O. trajectory will depend on the the initial guess in $\underline{b}_f$, but then for stable systems, the O.L.F.O. trajectory will not vary drastically when we vary the initial guess in $\underline{b}_f$.

(5) For the unstable system, the O.L.F.O. trajectory seems to follow closely its input-free trajectory in the beginning, until the diverging phenomenon tells the identifier to send back large controls for identification purposes. This causes some overshoots in the trajectory. The magnitude of the maximum overshoot seems to relate inversely with the values for the weighting constant h on control. For stable systems, simultaneous identification and control seem to be carried out in the beginning. Since the system is stable, with little control energy, the state will go to zero, so after some time period when the state is near the origin, approximately zero control is applied thus terminating the identification of $\underline{b}$.

(6) Lastly, the author would like to comment on the computational feasibility of the proposed scheme. The above experiments were simulated using an IBM 360/64/40 system. It was found that the actual computation of the O.L.F.O. control sequence can be carried out almost in real time for $N = 40$; i.e. in about 0.2 seconds, the following tasks were accomplished: One step computation of (6.3.19)-(6.3.32) (6 vector difference equation and 6×6 matrix difference equation), the parameter computations (6.3.34)-(6.3.37), and the computation of $\underline{\tilde{K}}(k/k)$ (6.3.32), $\underline{S}(k)$(6.3.39)

(one 12 × 12 matrix difference equation and one 3 × 3 matrix difference equation, computed in a time-backward direction directly for $k \leq 40$ steps, $k = 0, 1, \ldots, N-1$).

Further Experimental Studies

The following experiments are suggested so as to provide a deeper understanding on this class of problems.

(1) Implement a window-shifting O.L.F.O. control sequence as was suggested in section 6.7. This will allow us to consider control over an infinite time span for $k = 0, 1, \ldots$. To increase the convergent rate, apply control sequence

$$\tilde{u}^*(k) = \begin{cases} u^*(k) & \text{if} \quad u^*(k) \geq \varepsilon \\ \varepsilon & \text{if} \quad u^*(k) < \varepsilon \end{cases} \tag{7.24}$$

if $||\underline{\Sigma}_b^o(k/k)|| \geq \delta$, and $\tilde{u}^*(k) = u^*(k)$ if $||\underline{\Sigma}_b^o(k/k)|| < \delta$. The values for δ and ε are adjusted through experimentation.

(2) Design a computer program which will enable us to study the statistical behavior of the O.L.F.O. control system. For a fixed assumed structure of the system and the same weighting constants, study the statistical behavior of the system and the average convergence rate of the suboptimal control system to the optimal system. Vary the weighting constant h on the control, and investigate, in a statistical sense, how it affects the average maximum overshoot in the trajectory.

(3) To avoid large overshoots in the beginning for the unstable system, one may wish to have a large weighting factor h for the control energy in the beginning, and when the true value of

$\underline{b}$ is exactly recovered, we may want $h = 0.1$. Thus we may prefer h to be time varying

$$h(k) = g(k) + 0.1 \tag{7.25}$$

where $g(k)$ is nonincreasing and $g(k) \to 0$ as $k \to \infty$. Such an ad-hoc approach may lead to a well behaved O.L.F.O. control system.

(4) The assumption that $\underline{b} \sim G(\underline{b}_o, \underline{\Sigma}_{bo})$ is made for mathematical convenience. In actual practice, $\underline{b}_o$ and $\underline{\Sigma}_{bo}$ may not be available. With the results in chapter 3, observability of the pair $(\underline{A}, \underline{C})$ is sufficient to assure that independent of the guess on $\underline{b}_o$, asymptotic convergence of the estimate of $\underline{b}$ is obtained. But it would be important to find out how different assumptions on $\underline{b}_o$ and $\underline{\Sigma}_{bo}$ will effect the rate of convergence for both stable and unstable systems.

(5) By varying the sampling rate, one can study the effect of sampling period to the behavior of the overall suboptimal O.L.F.O. control system.

CHAPTER VIII

CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

The observer theory for discrete and continuous time linear systems have been developed in parallel. We showed that one can view an observer-estimator as a learning device which is used to learn all recoverable uncertainties while taking the statistical behavior of all inherent disturbances into consideration. The class of observers-estimators which will do the learning optimally in the mean square sense is also derived. Such optimal classes of observers-estimators can be incorporated in the overall optimal control system, and for this reason analytical studies on the optimum class of observers-estimators was carried out in detail. It is noted that observers theory includes Kalman filtering and deterministic exponential estimation as special cases.

The stochastic control of linear systems with known dynamics was treated in detail. For this class of problems, we have imperfect information due to the fact that there are inherent noise disturbance and unknown initial condition of the system being controlled. It was proved that for quadratic criteria the optimal controller consists of a "learner" and a set of feedback gains. The learner is realized by an optimum observer-estimator. The result is also known as the Separation Theorem. Physically, the operating function of the optimum observer-estimator is to learn the current state of the system. It can be shown that if the current state of the system is asymptotically recoverable and if the system can be stabilized by adapting some feedback gain, then the overall optimal stochastic system will have nice behavior. The approach taken in studying this specific class of problems can be extended to more general classes of problems where the cost criteria are other than quadratic.

In the next level, we considered control of linear discrete systems with unknown gain parameters. Since the truly optimal control sequence cannot be obtained because of the "curse of dimensionality." we look for a computationally feasible suboptimal control sequence. Prompted by physical consideration and computational considerations, we used the open-loop feedback optimal approach to derive the O.L.F.O. control sequence. It was proved that the O.L.F.O. controller consists of a learner, which we call an identifier, and a feedback gain plus correction term. The identifier is realized by an optimal observer-estimator whose operating function is to learn the current state and current unknown gain. Analytical studies were done on the overall O.L.F.O. control system. It was proved that if the initial state and unknown gain parameters are recoverable. then the overall O.L.F.O. control system will asymptotically converge to the truly optimal stochastic control system. The derived results seem to be computationally feasible. The computation of the O.L.F.O. control is done on-line. For all time $k = 0, 1, \ldots, N-1$, we have to compute a one-step 2n-vector difference equation and a one-step $2n \times 2n$ matrix difference equation (identification equations), then a $(N - k)$-steps n-vector difference equation and a $(N - k)$-steps $n \times n$ matrix difference equation (parameters computation), and finally a $(N - k)$-steps $(n+1)n \times (n+1)n$ matrix difference equation (computation of $\underline{\tilde{K}}(k/k)$(see Fig. 6.2 and Fig. 6.3). The vectors and matrices being stored as time unit advances are $\underline{\hat{x}}(k/k)$, $\underline{\hat{b}}(k/k)$, and $\underline{\Sigma}(k/k)$ which require a total of $(2n^2 + 3n)$ memory locations. (Note that $\underline{\Sigma}(k/k)$ is symmetric and this cuts down the storage memory requirements.)

Using the theoretical results derived, a computer program is developed to study the control of a variety of third order systems with known poles but unknown zeroes. Sample runs were made mainly to study the convergence

rate of the O.L.F.O. control system to the truly optimal system. The experimental results seem to indicate that the rate of convergence depends on the structure of the system: stable plants appear to have slow convergence, whereas unstable plants will result in fast convergence. For stable system, the convergence rate depends highly on the initial guess of the unknown zeroes locations; but for unstable stable, it appears that the rate of convergence is quite insensitive to the initial guess of the unknown zeroes locations. More experiments must be performed so as to obtain a deeper understanding on this class of problems and obtain engineering rules-of-thumb.

Directions of further research which are related directly to this work are suggested near the end of each chapter when appropriate. In the following, a list of topics is given, which the author thinks is a continuation of this present work. Some possible approaches to these different problems are suggested and the applicability of the results obtained in this thesis to these different problems is discussed.

(A) <u>Stochastic Control of Continuous-Time Linear Systems With Unknown Gains</u>

We consider a continuous analog of (6.2.1)

$$S^c: \quad \begin{aligned} &\dot{\underline{x}}(t) = \underline{A}(t)\underline{x}(t) + \underline{b}(t)u(t) + \underline{\xi}(t) \quad ; \quad \underline{x}(t_o) \sim \mathcal{G}(\underline{x}_c, \underline{\Sigma}_{xo}) \\ &\underline{y}(t) = \underline{C}(t)\underline{x}(t) + \underline{\eta}(t) \end{aligned} \tag{6.8.1}$$

the gain vector $\underline{b}(t)$ is unknown but satisfies the stochastic differential equation:

$$\dot{\underline{b}}(t) = \underline{G}(t)\underline{b}(t) + \underline{\gamma}(t) \quad : \quad \underline{b}(t_o) \sim \mathcal{G}(\underline{b}_o, \underline{\Sigma}_{bo}) \tag{6.8.2}$$

The noises, $\underline{\xi}(t)$, $\underline{\eta}(t)$, and $\underline{\gamma}(t)$ are assumed to be white Gaussian with known statistical law. The performance measure is

$$J^c(u) = E\left\{\underline{x}'(T)\ \underline{F}\ \underline{x}(T) + \int_{t_o}^{T} \left[\underline{x}'(t)\ \underline{W}(t)\ \underline{x}(t) + u^2(t)h(t)\right]dt\right\} \qquad (6.8.3)$$

The control problem is to find $u(\tau)$, $\tau \in [t_o, T)$, such that (6.8.3) is minimized subject to dynamic constraints, (6.8.1) and (6.8.2). Instead of first taking a sample data version of the problem and then applying the derived results in chapter 6 (see chapter 7), we can apply the O.L.F.O. approach to the continuous time system directly. One would then obtain a continuous time identifier which estimates the current state and current gain in continuous time. The results in chapter 4 can be applied. As analogous to the discrete time version, we would then formulate a deterministic (continuous-time) open-loop control problem. One may expect the overall O.L.F.O. control system in the continuous-time case will be similar in structure to that in the discrete version. The main difficulty lies in the capability of computing the O.L.F.O. adaptive gain and the correction term in <u>continuous</u> time. Some modifications can be made which take computation capability into account. One approach may be that we resolve the open-loop problem only in discrete time, $t = 0, \Delta, 2\Delta, 3\Delta, \ldots$, even though we have continuous time observation.

(B) <u>Control With Unknown Dynamics</u>

Consider the problem of controlling an unknown system $\mathcal{S}$, (6.2.1), where the matrix $\underline{A}(k)$, $k = 0, 1, 2, \ldots$, is unknown but satisfies some linear difference equation. The statistical laws of the noise are assumed known. Our objective is to control the system $\mathcal{S}$ using the quadratic criteria. Formally, the truly optimal control can be obtained if we can solve Bellman's equation. Unfortunately, this is impossible with the present stage of development of computer technology. Therefore, one can look for suboptimal but computationally feasible solutions to the problem.

It would be desirable if we can have analytical studies on the derived suboptimal control system. Different approaches guided by engineering intuition are possible. An approach, which is a combination of maximum likelihood and O.L.F.O. is suggested where an analytic study of the behavior of the overall suboptimal system may be possible.

Consider the augmented system $\tilde{S}$ given by (6.3.24). Let $U(0, k-1)$ be applied and $Y_{U(0,k-1)}(0,k)$ is observed. The most probable estimate (maximum likelihood estimate) of $\{\underline{A}(i)\}_{i=0}^{k-1}$, which is denoted by $\{\underline{A}_k^o(i)\}_{i=0}^{k-1}$, is obtained by picking $\{\hat{\underline{A}}(i)\}_{i=0}^{k-1}$ to maximize the conditional probability density $p\left(\{\hat{\underline{A}}(i)\}_{i=0}^{k-1} \middle| Y_{U(0,k-1)}(0,k)\right)$ subject to a certain difference equation describing the evoluation of $\underline{A}(i)$, $i = 0, 1, \ldots, k-1$. Extrapolate the estimate of $\{\underline{A}(i)\}_{i=k}^{\infty}$ and the estimates are denoted by $\{\underline{A}_k^o(i)\}_{i=k}^{\infty}$. Assume that $\{\underline{A}_k^o(i)\}_{i=0}^{\infty}$ is the true $\underline{A}(i)$, $i = 0, 1, \ldots$, and apply the results of chapter 6. The whole procedure is repeated at every step, $k = 0, 1, \ldots$.

Theoretically, this approach has some advantageous features. Using Wald's Theorem [68], one will obtain asymptotic consistent (with probability 1) estimate of $\{\underline{A}(i)\}_{i=0}^{\infty}$; one can then apply the results of section 6.6 to obtain overall asymptotic optimal control system.

The difficulty lies in the real time computation of $\{\underline{A}_k^o(i)\}_{i=0}^{\infty}$, $k = 0, 1, \ldots$ using a computer. For references in maximum likelihood estimation, see Kashyap [67], Wald [68], Rauch, Tung, and Striebel [44]; for evaluation of likelihood functions of a Gaussian process, see also Schweppe [69].

(C) Control With Unknown Gain and Imperfectly Known Disturbance

Assume that the matrices $\underline{A}(k)$, $k = 0, 1, \ldots$, are known, the gain vectors $\underline{b}(k)$, $k = 0, 1, \ldots$, are assumed to be unknown but described by

(6.2.2), (6.2.4), and (6.2.7). The vectors $\underline{n}(k)$, $\underline{\xi}(k)$, $\underline{\gamma}(k)$, $k = 0, 1, \ldots.$ are independent Gaussian vectors with unknown means and/or covariances. It is necessary for us to recover the true means and covariances of the noise vectors. A combination of maximum likelihood and O.L.F.O. approach can be applied to such class of problems.

For references which are related to this class of problems, see Saga and Husa [70], Taran [71], Kashyap [67].

With some thorough understanding in the problems (A) and (B), we can then start to investigate the problem of controlling a system where $\underline{A}(k)$, $\underline{b}(k)$, $k = 0, 1, \ldots.$ are unknown but satisfy some difference equations, and the noise vectors are independent Gaussian vectors with unknown means and variances.

APPENDIX A

ON THE PSEUDO-INVERSE OF A MATRIX

Let $\underline{A}$ be an $n \times m$ matrix which maps $R^m \to R^n$. The pseudo-inverse of $\underline{A}$ is denoted by $\underline{A}^{\#}$ and satisfies the conditions:

(1) $\underline{A}^{\#}\underline{Ax} = \underline{x} \quad ; \quad \forall \ \underline{x} \in R_a(\underline{A}')$ (A.1)

(2) $\underline{A}^{\#}\underline{z} = \underline{0} \quad ; \quad \forall \ \underline{z} \in N(\underline{A}')$ (A.2)

(3) $\underline{A}^{\#}(\underline{y} + \underline{z}) = \underline{A}^{\#}\underline{y} + \underline{A}^{\#}\underline{z} \quad ; \quad \forall \ \underline{y} \in R_a(\underline{A}), \ \underline{z} \in N(\underline{A}')$ (A.3)

With this definition, we have the following properties:

(A) $(\underline{A}^{\#})^{\#} = \underline{A}$ (A.4)

(B) $\underline{A}^{\#} \ \underline{A} \ \underline{A}^{\#} = \underline{A}^{\#}$ (A.5)

(C) $\underline{A} \ \underline{A}^{\#} \ \underline{A} = \underline{A}$ (A.6)

(E) Let $\underline{A}$ be an $n \times m$ matrix $(n \geq m)$ of rank m. Then

$$\underline{A}^{\#} = (\underline{A}' \ \underline{A})^{-1} \ \underline{A}' \tag{A.7}$$

For the proofs of (A.4)-(A.7), see Zadah and Desoer [48], Levine [23]; for a different approach to generalized inverse of a matrix, see Penrose [72].

APPENDIX B

WEINER-HOPF EQUATION

Let $F(k) \triangleq F(y(i);\ i = 0, 1, \ldots, k)$, we have $F(i) \subset F(i+1)$, $i = 0, 1, \ldots, k-1$, and so $\underline{y}(i)$ is $F(k)$-measurable for $i = 0, 1, \ldots, k$. Using lemma 2.2.6, and lemma 2.2.7, we have

$$E\{\underline{x}(k)\ \underline{y}'(i)\} = E\left\{E\{\underline{x}(k)\ \underline{y}'(i)/F(k)\}\right\} = E\{\hat{\underline{x}}(k/k)\ \underline{y}'(i)\} \qquad i = 1, \ldots, k \tag{B.1}$$

By assumption, $\underline{w}(k)$ satisfies (3.3.18), thus

$$E\{(\underline{w}(k) - \hat{\underline{x}}(k/k))\underline{y}'(i)\} = \underline{0} \quad i = 0, 1, \ldots, k\ ;\ k = 0, 1, \ldots. \tag{B.2}$$

Since both $\underline{w}(k)$, $\hat{\underline{x}}(k/k)$ are linear functioals of $\underline{y}(0), \ldots, \underline{y}(k)$, (B.2) also implies

$$E\{(\underline{w}(k) - \hat{\underline{x}}(k/k))(\underline{w}(k) - \hat{\underline{x}}(k/k))'\} = \underline{0} \quad k = 0, 1, \ldots. \tag{B.3}$$

Thus $\underline{w}(k) = \hat{\underline{x}}(k/k)$ a.s.

The proof of Weiner-Hopf equation for the continuous case is the same with slight modification, the induced σ-algebra $F(k)$ is replaced by $F_t \triangleq F\{\underline{y}_1(\tau),\ \tau \in [t_o, t),\ \underline{y}_2(\tau),\ \tau \in [t_o, t]\}$. And so if $\underline{w}(t);\ t \geq t_o$ is a random process such that for $t \geq t_o$, $\underline{w}(t)$ is a linear functional of $\underline{y}_1(\tau),\ \tau \in [t_o, t)$, and $\underline{y}_2(\tau),\ \tau \in [t_o, t]$; and $\underline{w}(t)$ satisfies

$$E\{\underline{w}(t)\ \underline{y}_1'(\tau)\} = \{E\ \underline{x}(t)\ \underline{y}_1'(\tau)\} \qquad \tau \in [t_o, t)\ ;\ t \geq t_o \tag{B.4}$$

$$E\{\underline{w}(t)\ \underline{y}_2'(\tau)\} = \{E\ \underline{x}(t)\ \underline{y}_2'(\tau)\} \qquad \tau \in [t_o, t]\ ;\ t \geq t_o \tag{B.5}$$

then $\underline{w}(t) = \hat{\underline{x}}(t/t)$ a.s., $t \geq t_o$. (B.4), (B.5) imply the projection equations (4.3.44).

APPENDIX C

EQUATION FOR ERROR PROCESS (CONTINUOUS TIME CASE)

Let $\underline{x}(t)$ be a random process given by (4.3.1), and $\underline{w}(t)$ be a random process satisfying (4.3.19), (4.3.20), and (4.3.16). Define $\underline{e}(t) \triangleq \underline{w}(t) - \underline{x}(t)$. Differentiating $\underline{e}(t)$ and using (4.3.19), (4.3.1), and (4.3.12):

$$
\begin{aligned}
\underline{e}(t) = {} & \dot{\underline{P}}(t)\, \underline{z}(t) + \underline{P}(t)\, \underline{T}(t)\, \underline{A}(t)\, \underline{w}(t) - \underline{P}(t)\, \underline{T}(t)\, \underline{L}_1(t)\, \underline{C}_1(t)\, \underline{w}(t) \\
& + \underline{P}(t)\, \dot{\underline{T}}(t)\, \underline{P}(t)\, \underline{z}(t) + \underline{P}(t)\, \dot{\underline{T}}(t)\, \underline{V}_2(t)\, \underline{C}_2(t)\, \underline{x}(t) \\
& + \underline{P}(t)\, \underline{T}(t)\, \underline{L}_1(t)\, \underline{C}_1(t)\, \underline{x}(t) + \underline{P}(t)\, \underline{T}(t)\, \underline{L}_1(t)\, \underline{n}(t) \\
& + \underline{P}(t)\, \underline{T}(t)\, \underline{B}(t)\, \underline{u}(t) + \dot{\underline{V}}_2(t)\, \underline{C}_2(t)\, \underline{x}(t) + \underline{V}_2(t)\, \dot{\underline{C}}_2(t)\, \underline{x}(t) \\
& + \underline{V}_2(t)\, \underline{C}_2(t)\, \underline{A}(t)\, \underline{x}(t) + \underline{V}_2(t)\, \underline{C}_2(t)\, \underline{B}(t)\, \underline{u}(t) + \underline{V}_2(t)\, \underline{C}_2(t)\, \underline{\xi}(t) \\
& - \underline{A}(t)\, \underline{x}(t) - \underline{B}(t)\, \underline{u}(t) - \underline{\xi}(t) \\
= {} & \left(\dot{\underline{P}}(t) + \underline{P}(t)\, \dot{\underline{T}}(t)\, \underline{P}(t)\right) \underline{z}(t) - \left(\dot{\underline{P}}(t)\, \underline{T}(t) + \underline{P}(t)\, \dot{\underline{T}}(t)\, \underline{P}(t)\, \underline{T}(t)\right) \\
& \cdot \underline{x}(t) + \underline{P}(t)\, \underline{T}(t)\, \underline{A}(t)\, \underline{e}(t) - \underline{P}(t)\, \underline{T}(t)\, \underline{L}_1(t)\, \underline{C}_1(t)\, \underline{e}(t) \\
& + \underline{P}(t)\, \underline{T}(t)\, \underline{L}_1(t)\, \underline{n}(t) - \underline{P}(t)\, \underline{T}(t)\, \underline{\xi}(t)
\end{aligned}
\tag{C.1}
$$

Since $\underline{P}(t)$, $T(t)$, $V_2(t)$, $\underline{C}_2(t)$ also satisfy (4.3.16), we have

$$
\dot{\underline{T}}(t)\, \underline{P}(t) + \underline{T}(t)\, \dot{\underline{P}}(t) = \underline{0} \quad ; \quad \underline{C}_2(t)\, \dot{\underline{P}}(t) + \dot{\underline{C}}_2(t)\, \underline{P}(t) = \underline{0}
\tag{C.2}
$$

and so we have

$$
\begin{aligned}
\dot{\underline{P}}(t) + \underline{P}(t)\, \dot{\underline{T}}(t)\, \underline{P}(t) &= \left(\underline{I}_n - \underline{P}(t)\, \underline{T}(t)\right) \dot{\underline{P}}(t) = \underline{V}_2(t)\, \underline{C}_2(t)\, \dot{\underline{P}}(t) \\
&= -\, \underline{V}_2(t)\, \dot{\underline{C}}_2(t)\, \underline{P}(t)
\end{aligned}
\tag{C.3}
$$

$$
\begin{aligned}
\dot{\underline{P}}(t)\, \underline{T}(t) + \underline{P}(t)\, \dot{\underline{T}}(t)\, \underline{P}(t)\, \underline{T}(t) &= -\, \underline{V}_2(t)\, \dot{\underline{C}}_2(t)\, \underline{P}(t)\, \underline{T}(t) \\
&= -\, \underline{V}_2(t)\, \dot{\underline{C}}_2(t) + \underline{V}_2(t)\, \dot{\underline{C}}_2(t)\, \underline{V}_2(t)\, \underline{C}_2(t)
\end{aligned}
\tag{C.4}
$$

Substituting (C.3) and (C.4) into (C.1) and noting $\underline{w}(t) = \underline{P}(t)\ \underline{z}(t) + \underline{V}_2(t)\ \underline{y}_2(t)$, we have

$$\dot{\underline{e}}(t) = (\underline{A}(t) - \underline{V}_2(t)\ \tilde{\underline{C}}_2(t) - \underline{P}(t)\ \underline{T}(t)\ \underline{L}_1(t)\ \underline{C}_1(t))\ \underline{e}(t) + (\underline{V}_2(t)\ \underline{C}_2(t) - \underline{I}_n)\ \underline{\xi}(t) + \underline{P}(t)\ \underline{T}(t)\ \underline{L}_1(t)\ \underline{n}(t) \tag{C.5}$$

The initial error is

$$\begin{aligned}\underline{e}(t_o) &= \underline{P}(t_o)\ \underline{z}(t_o) + \underline{V}_2(t_o)\ \underline{C}_2(t_o)\big(\underline{x}(t_o) - \underline{x}(t_o)\big)\\ &= \underline{P}(t_o)\ \underline{T}(t_o)(\underline{x}_o - \underline{x}(t_o)) = (\underline{I}_n - \underline{V}_2(t_o)\ \underline{C}_2(t_o))(\underline{x}_o - \underline{x}(t_o)) \end{aligned}\tag{C.6}$$

BIBLIOGRAPHY

1. Doob, J. L., Stochastic Processes, Wiley and Sons, Inc., New York, 1953.

2. Loève, M. Probability Theory, Van Nostrand, New York, 3rd edition, 1963.

3. Davenport, W. B., Jr., and Root, W. L., Random Signals and Noise, McGraw-Hill Book Co., Inc., 1958.

4. Wonham, W. M., Lecture Notes on Stochastic Control, Lecture Notes 67-2, Center for Dynamical Systems, Division of Applied Mathematics, Brown University, February 1967.

5. Tse, E., "Generalized Random Processes," Report ESL-R-366, M.I.T., October 1968.

6. Tricomi, F. G., Integral Equations, New York, Interscience Publishers, 1957.

7. Dynkin, E. B., Markov Process, Academic Press, New York, 1965.

8. Skorokhod, A. V., Studies in the Theory of Random Process, Addison-Welsey, New York, 1965.

9. Tse, E., "Application of Pontryagin's Minimum Principles to Filtering Problems," M.S. Thesis, E.E., M.I.T., June 1967.

10. Athans, M., and Tse, E., "A Direct Derivation of the Optimal Linear Filter Using the Maximum Principle," IEEE Trans. on Automatic Control, December 1967.

11. Pontryagin, L. S., Boltyanskii, V. G., Gamkrelidze, R. V., and Mishchenko, E. F., The Mathematical Theory to Optimal Processes, New York, Interscience, 1962.

12. Athans, M., "The Matrix Minimum Principle," Information and Control, Vol. 11, December 1967.

13. Halmas, P., and Savage, L. J., "Application of the Radon-Nikodym Theorem of Sufficient Statistics," Ann. Math. Stat. 20, 1949.

14. Meyer, P. A., Probability and Potentials, Blaisdell Publishing Company, 1966.

15. Cramer, H., Mathematical Methods of Statistics, Princeton University Press, Princeton, New Jersey, 1946.

16. Mckean, H. P., Stochastic Integral, Academic Press, 1969.

17. Ito, K., "On Stochastic Differential Equations," Mem. Amer. Math. Soc., 4, 1951.

18. Gelfand, I. M., and Vilenkin, N. Y., Generalized Functions, Vol. 4, Academic Press, 1964.

19. Stratonovich, R. L., "A New Representation for Stochastic Integrals and Equations," SIAM J. Control 4(2), 1966.

20. Wong, E., and Zakai, M., "On the Relation Between Ordinary and Stochastic Differential Equations," Int. J. Engrg. Sci., 3, 1965.

21. Clark, J. M. C., "The Representation of Non-Linear Stochastic Systems with Applications to Filtering," Ph.D. Thesis, Electrical Engineering Department, Imperial College, London, 1966.

22. Wonham, W. M., "Random Differential Equations in Control Theory," to appear: A. T. Bharucha-Reid, Ed., "Probabilistic Methods in Applied Mathematics," Vol. II, Academic Press, 1969.

23. Levine, W. S., "Optimal Output-Feedback Controllers for Linear Systems," Ph.D. Thesis, M.I.T., January 1969.

24. Debs, A., "Optimal and Suboptimal Feedback Control for Deterministic Systems," Ph.D. Thesis, E.E. Department, M.I.T., February 1969.

25. Bellman, R., Dynamic Programming, Princeton University Press, Princeton, New Jersey, 1957.

26. Dreyfus, S. E., Dynamic Programming and the Calculus of Variations, Mathematics in Science and Engineering, Academic Press, 1965.

27. Wonham, W. M., "Stochastic Problems in Optimal Control," RIAS Tech. Rpt. 63-14, May 1963.

28. Deyst, J. J., Jr., and Price, C. F., "Conditions for Asymptotic Stability of the Discrete, Minimum-Variance, Linear Estimator," IEEE Trans. on Automatic Control, Vol. AC-13, No. 6, December 1968.

29. Lorenson, R. W., "On the Error Behavior in Linear Minimum Variance Estimation Problems," IEEE Trans. on Automatic Control, Vol. AC-12, No. 5, October 1967.

30. Aoki, M., Optimization of Stochastic Systems, Mathematics in Science and Engineering, Vol. 32, Academic Press, 1967.

31. Kleinman, D. L.,"On the Linear Regulator Problem and the Matrix Riccati Equation," ESL Report 271, M.I.T., 1966.

32. Wonham, W. M., "On a Matrix Riccati Equation of Stochastic Control," SIAM J. on Control, Vol. 6, No. 4, November 1968.

33. Holtzman, J. M., and Halkin, H., "Directional Convexity and Maximum Principle for Discrete System," J. SIAM Control, 4, 1966.

34. Bellman, R., Introduction to Matrix Analysis, McGraw Hill Book Co., Inc., New York, 1960.

35. Luenberger, D. G., "Observers for Multivariable Systems," IEEE Trans. on Automatic Control, Vol. AC-11, April 1966.

36. Luenberger, D. G., "Observing the State of a Linear System," IEEE Trans. on Military Electronics, Vol. MIL-8, April 1964.

37. Aoki, M., and Huddle, J. R., "Estimation of the State Vector of a Linear Stochastic System with Constrained Estimator," IEEE Trans. on Automatic Control, AC-12, August 1967.

38. Tse, E., and Athans, M., "Optimal Minimal-Order Observers-Filters for Discrete Linear Time-Varying Systems," Submitted to IEEE Trans. on Automatic Control.

39. Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems," Trans. ASME, J. Basic Engrg., Ser. D, Vol. 82, March 1960.

40. Kalman, R. E., "Fundamental Study of Adaptive Control Systems," Tech. Rept. ASK-TR-61, Vol. 1, App. I.

41. Kalman, R. E., "On the General Theory of Control Systems," Proc. First International Congress on Automatic Control, Moscow, 1960.

42. Athans, M., Wishner, R. P., and Bertolini, A., "Suboptimal State Estimation for Continuous-Time Nonlinear Systems from Discrete Noisy Measurements," IEEE Trans. on Automatic Control, Vol. AC-13, October 1968.

43. Bryson, A., Jr., and Ho, Y. C. Applied Optimal Control, Blaisdell Publishing Company, 1969.

44. Rauch, H., Tung, G., and Streibel, "On the Maximum Likelihood Estimates for Linear Dynamic Systems," AIAA J., Vol. 3, 1965.

45. Freeman, H., Discrete-Time Systems, John Wiley and Sons, Inc., 1965.

46. Henrikson, L. J., "Sequentially Correlated Measurement Noise with Application to Inertial Navigation," Spacecraft and Rockets, Vol. 5, 1968.

47. Van Trees, H. L., Detection, Estimation, and Modulation Theory, Part I, John Wiley and Sons, Inc., New York, 1968.

48. Zadeh, L. A., and Desoer, C. A., "Linear System Theory," McGraw Hill Book Company, New York, 1963.

49. Johnson, G. W., "A Deterministic Theory of Estimation and Control," JACC, 1969.

50. Kalman, R. E., and Bucy, R. S., "New Results in Linear Filtering and Prediction Theory," Trans. ASME, J. Basic Engrg., Vol. 33, 1961.

51. Weiner, N., Extrapolation, Interpolation and Smoothing of Stationary Time Series, John Wiley and Sons, New York, 1949.

52. Bucy, R. S., "Optimal Filtering for Correlated Noise," J. of Math. Analysis and Application, Vol. 20, No. 1, October 1967.

53. Mehra, R. K., and Bryson, A. E., Jr., "Linear Smoothing Using Measurements Containing Correlated Noise with an Application to Interial Navigation," IEEE Trans. on Automatic Control, AC-13, October 1968.

54. Root, W. L., "Singular Gaussian Measures in Detection Theory," Time Series Analysis, M. Rosenblatt, Ed., New York, J. W. Wiley, 1963.

55. Geesey, R., and Kailath, T., "An Innovation Approach to Least-Squares Estimation--Part III: Estimation in Colored Noise," to be published.

56. Joseph, P. D., and Tou, J. T., "On Linear Control Theory," AIEE Trans. (Appl. and Industry), September 1961.

57. Athans, M., and Falb, P., Optimal Control, McGraw Hill, New York, 1966.

58. Gunckel, T. L., and Franklin, G. F., "A General Solution for Linear, Sampled-Data Control," Trans. ASME, 85D, June 1963.

59. Streibel, C., "Sufficient Statistics in the Optimal Control of Stochastic Systems," J. of Math. Analysis and Appl., 12, 1965.

60. Farison, J. B., Graham, R. E., and Shelton, R. C., "Identification and Control of Linear Discrete Systems," IEEE Trans. on Automatic Control, AC-12, August 1967.

61. Murphy, W. J., "Optimal Stochastic Control of Discrete Linear Systems with Unknown Gain," IEEE Trans. on Automatic Control, AC-13, August 1968.

62. Gorman, D., and Zaboroszky, J., "Stochastic Optimal Control of Continuous Time Systems with Unknown Gain," IEEE Trans. on Automatic Control, AC-13, December 1968.

63. Bar-Shalom, Y., and Sivan, R., "On the Optimal Control of Discrete-Time Linear Systems with Random Parameters," IEEE Trans. on Automatic Control, AC-14, February 1969.

64 Florentin, J. J., "Optimal, Probing, Adaptive Control of a Simple Baysian System," J. Electronics and Control, Vol. 11, 1962.

65. Witsenhausen, H. S., "Minimax Controls of Uncertain Systems," Ph.D. Thesis, Department of Electrical Engineering, M.I.T., May 1966.

66. Schweppe, F. C., Notes for the course 6.606 offered at M.I.T., 1969.

67. Kashyap, R. L., "Maximum Likelihood Identification of Stochastic Linear Systems," to be published.

68. Wald, A., "Note on the Consistency of the Maximum Likelihood Estimate," Ann. Math. Stat., Vol. 20, No. 4, December 1949.

69. Schweppe, F. C., "Evaluation of Likelihood Functions for Gaussian Signals," IEEE Trans. on Information Theory, IT-11, January 1965.

70. Saga, A. P., and Husa, G. W., "Adaptive Filtering with Unknown Prior Statistics," JACC, 1969.

71. Taran, T. J., "Stochastic Optimal Control with Noisy Observation having Imperfectly Known Disturbances," JACC, 1969.

72. Penrose, R., "A Generalized Inverse for Matrices," Proc. Cambridge Phil. Soc., 51, 1955.

73. Fel'dbaum, A. A., Optimal Control Systems, Mathematics in Science and Engineering, Vol. 22, Academic Press, 1965.

74. Kushner, H. J., Stochastic Stability and Control, Mathematics in Science and Engineering.

75. Levis, A., "On the Optimal Sampled-Data Control of Linear Processes," Ph.D. Thesis, M.I.T., June 1968.

Security Classification

DOCUMENT CONTROL DATA - R & D

1. ORIGINATING ACTIVITY (Corporate author) Massachusetts Institute of Technology Department of Electrical Engineering Cambridge, Massachusetts 02139	UNCLASSIFIED
3. REPORT TITLE ON THE OPTIMAL CONTROL OF LINEAR SYSTEMS WITH INCOMPLETE INFORMATION	
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Scientific Interim	
5. AUTHOR(S) (First name, middle initial, last name) Edison Tack-Shuen Tse	
6. REPORT DATE January 1970	7a. TOTAL NO. OF PAGES 285 7b. NO. OF REFS 75
8a. CONTRACT OR GRANT NO. AFOSR-69-1724 b. PROJECT NO. 9749-01 c. 61102F 681304 d.	9a. ORIGINATOR'S REPORT NUMBER(S) ESL-R-412 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) AFOSR 70-0474TR
10. DISTRIBUTION STATEMENT 1. This document has been approved for public release and sale; its distribution is unlimited	
11. SUPPLEMENTARY NOTES TECH, OTHER	12. SPONSORING MILITARY ACTIVITY Air Force Office of Scientific Research (SRMA) 1400 Wilson Boulevard Arlington, Virginia 22209

13. ABSTRACT

The control of linear systems with incomplete information is considered where the unknown disturbances and/or random parameters are assumed to satisfy some statistical laws.

The observer theory for linear systems is developed which generalizes the concepts due to Kalman and Luenberger pertaining to the design of linear systems which estimate the state of a linear plant on the basis of both noise-free and noisy measurements of the output variables. The Separation Theorem for linear system is then extended for such observers-estimators.

The problem of controlling a linear system with unknown gain is then considered. An open-loop-feedback-optimal control algorithm is developed which seems to be computationally feasible. Existence of such suboptimal control scheme is proved under the assumption that the unvertainties in the unknown gain will not grow in time. Covergence of such suboptimal control system to the truly optimal control system is considered. A computer program is developed to study the control of a variety third order systems with known poles but unknown zeroes. The experimental results serve to provide us with some more insights into the structure and behavior of the open-loop-feedback-optimal control systems.

DD FORM 1 NOV 65 1473

Security Classification